The Journal of Development Studies, 2018
Vol. 54, No. 2, 256–270, https://doi.org/10.1080/00220388.2017.1283014

Agricultural Commercialisation and Food Security in Rural Economies: Malawian Experience

NATALIA RADCHENKO* & PAUL CORRAL**
*Department of Economics, American University, Washington, DC, USA, **Poverty and Equity Global Practice, World Bank, Washington, DC, US

(Original version submitted August 2016; final version accepted January 2017)

ABSTRACT *This paper contributes to the debate on the nutrition-related outcomes of cash crop adoption by using a model with essential heterogeneity and a semi-parametric estimation technique. The model explicitly frames non-separability between production and consumption decisions of farming households providing an original test of separability. The empirical application is run using Malawian data. The results imply rational anticipations and decision process of agrarian households relative to the crop portfolio choice, disparate strength of market barriers faced by the farmers, non-separability between production and consumption decisions and a weak transmission from agricultural incomes to higher food expenditures and better diet.*

1. Introduction

The debate over the effects of agricultural commercialisation, despite its potential for increased household incomes in low-income agrarian economies, is centred over food security concerns. On the one hand, agricultural commercialisation is deemed to increase food security and improve household nutrition through increased revenue providing the necessary cash to buy farming inputs and marketed food. The alternative view associates food security with self-sufficiency, implying multiple structural constraints and output market failures that limit households' ability to utilise cash income in order to enhance food production and consumption. The constraints regard a highly underdeveloped private sector and poor rural infrastructure that yield fragmented and unreliable product and factor markets, underdeveloped input and output distribution networks, underserved remote rural areas whose holdings are inaccessible to traders, high transaction costs, and unavailable transport and credit for farmers and traders (Chinsinga, 2011; Harashima, 2008).

Carletto, Ruel, Winters, and Zezza (2015) summarise the main concepts pertaining to the link between agriculture and nutrition and review recent empirical findings. They point to the challenging task of identifying the causal relationship between agriculture and nutrition and the inconclusiveness of empirical studies that range from reporting zero or negative nutritional effects from agricultural commercialisation (Von Braun Kennedy 1994 provide an overview of 1990s studies and report findings from Gambia, Guatemala, Kenya, Malawi, the Philippines, and Rwanda; more recently, Masanjala (2006) and Goldstein et al. (2013) use Malawian and Benin data respectively) to low or high positive elasticity of nutrition-related outcomes with respect to agricultural income and diversification (see for example Strauss and Thomas (1995) or the more recent Nigeria, Zambia and Tanzania studies of Dillon, McGee, and Oseni (2015), Kumar, Harris, and Rawat (2015), and Slavchevska

Correspondence Address: Natalia Radchenko, American University, 4400 Massachusetts Ave NW, Kreeger 104, Washington, DC, 20016-8029, USA Email: nataliar@american.edu.

(2015) respectively). Yet, the majority of studies focus on agricultural income effect rather than on agricultural commercialisation impacts specifically. Summarising the literature, Carletto et al. (2015) suggest that the relationship is deeply complex with heterogeneous households' responses and non-linear elasticity of nutritional outcomes.

Employing the Malawian Third Integrated Household Survey (IHS3) and a model with essential heterogeneity (Heckman, Urzua, & Vytlacil, 2006), this paper renews the debate by focusing on the relationship between agricultural commercialisation and nutrition-related outcomes of farming households. Following the literature (Carletto et al., 2015, 2016), we consider different kinds of outcomes. First, household food expenditures represent marketed food consumption related to cash cropping: agricultural commercialisation displaces staple crops by cash crops (as displacing maize by tobacco in the Malawian context) but is expected to increase households' purchasing power; higher food expenditures under cash cropping regime as compared to food only cropping regime imply an opportunity to commercialise cash crops and/or buy market food; this means operational markets providing food security and eventual nutritional improvement. On the other hand, lower food expenditures under cash cropping may imply worsening of nutrition since household's food production is reduced in this case without being compensated by the purchase of market food.

While possibly improving food security, higher food expenditures do not necessarily mean improved nutrition in terms of a sufficiently varied diet (for example, malnutrition related to over-reliance on maize is one of the main concerns in Malawi). The second outcome used, the food consumption score, represents therefore, an intensive margin of food consumption: higher food expenditures under cash cropping could possibly not lead to a higher amount of food consumed, but to a more varied diet of the household. Once again, more varied diet under cash cropping as associated with a higher food consumption score, means marketing possibilities which allow for varied/more expensive food consumption bundles.

The contribution of the paper consists of modelling heterogeneous differentials of the household's nutrition-related outcomes under different cropping regimes and allowing for correlation between these differentials and the likelihood of engaging in agricultural commercialisation. These features extend the literature in several ways: first, they allow for addressing the impact of agricultural commercialisation on the household's food consumption in a more flexible way by assuming disparate impacts across the rural population; second, they allow for detecting different intensity of the barriers to selling agricultural output and/or buying marketed food using the cash revenues which are faced by the agrarian population due to market failures; third, they permit testing of farmers' anticipation and responses to disparate intensity of the barriers and examine the extent to which food market failures affect farmers' decisions relative to cash crop adoption.

Finally, by allowing for correlation between the differentials of the nutrition-related outcomes and the likelihood of agricultural commercialisation, the approach explicitly frames non-separability between production and consumption decisions of the farming households and provides an original empirical test of separability. The results support non-separability between production and consumption decisions of Malawian households and point out the dissimilar strength of market barriers faced by farmers.

The paper is organised as follows. Section 2 describes the Malawian data and background. Sections 3 and 4 outline the main concepts and model outcomes respectively. Section 5 discusses the results. Section 6 concludes.

2. Malawian data and background

The data used are taken from the Third Integrated Household Survey (IHS3) administrated in Malawi during 2010–2011 by the Malawi National Statistics Office. The IHS3 is representative at the national, urban, rural, and regional levels and provides household and individual characteristics. The data also contain information on assets, households' agricultural inputs and outcomes as well as community features. The IHS3 sample size is 12,271 households coming from 768 enumeration areas. The sub-sample used in the study consists of 9514 rural households with non-missing information on the key variables (the summary statistics on the key variables are provided in Table 1). Two distinct groups are defined among farming households: cash crop adopters (1569 households that plant either cotton or tobacco on their land along with other edible crops[1]) and growers of only food items (7945 households).

The Malawian economy is mainly agrarian with rural households making up about 90 per cent of the population or 10,000 out of 12,271 households in terms of the IHS3 survey. A large share of farmers are extremely poor, have very small landholdings (according to the IHS3, the average amount of land owned by farming households is about one hectare) and little wealth. In order to survive, and with a lack of functioning input and output markets, uncertainty about input and output prices, high

Table 1. Summary statistics, by crop portfolios

	All		Only food growers		Cash crop adopters		
Variable Name	Mean	SE	Mean	SE	Mean	SE	
Ln of annual per capita food expenditures, (ln USD)	4.18	(0.64)	4.17	(0.64)	4.24	(0.62)	***
Ln of food consumption score (FCS)[a]	4.19	(0.26)	4.19	(0.26)	4.21	(0.24)	**
Harvest value [b], (ln USD)	4.86	(1.12)	4.65	(1.03)	5.95	(0.93)	***
Acreage, Ha	0.73	(0.63)	0.65	(0.57)	1.18	(0.73)	***
Household agro-ecological conditions							
Potential wetness index [c]	13.5	(2.34)	13.49	(2.40)	13.57	(2.02)	
Workability constraint index [d]	0.49		0.47		0.62		***
Tropic warm/semiarid	0.48		0.47		0.55		***
Tropic-warm/subhumid	0.35		0.37		0.19		***
Tropic-cool/semiarid	0.1		0.09		0.18		***
Tropic-cool/subhumid	0.07		0.07		0.08		**
Demographic characteristics							
Head's age	43	(16.00)	43	(17.00)	41	(14.00)	***
Household size	4.7	(2.18)	4.59	(2.17)	5.23	(2.14)	***
Child dependency ratio [e]	0.77	(0.70)	0.77	(0.72)	0.79	(0.62)	
Female headed household	0.25		0.28		0.1		***
Head's education							
None	0.76		0.77		0.72		***
Primary	0.1		0.09		0.15		***
Secondary	0.13		0.13		0.12		
Tertiary	0.01		0.01		0.01		**
Community feature Indexes							
Manufactured goods access improved? [f]	0.73		0.71		0.79		***
Staple foods access stable?	0.08		0.08		0.11		***
Distances							
Distance to population center, Km	38.11	(21.43)	37.06	(21.86)	43.68	(17.99)	***
Distance to border, Km	24.91	(18.24)	23.86	(18.29)	30.51	(16.90)	***
Distance to tobacco auction floor, Km	76.87	(46.22)	76.64	(47.16)	78.15	(40.87)	
Number of observations	9,514		7,945		1,569		

*Notes:****, **, * indicate significance of the difference of means at the 1, 5, and 10 per cent level respectively. [a] A composite score of intensity and dietary diversity based on consumption of staple grains and tubers, pulses, vegetables, fruits, meat and fish, dairy, oil and sugar in the past seven days.
[b] The aggregate of the kilograms of each crop harvested multiplied by the median price per kilogram of the crop within the enumeration area. [c] A measure of topographic wetness conditions provided by SRTM (Shuttle Radar Topography Mission) 90m Digital Elevation Database and based on flow accumulation or effective drainage area and the area topography; it not only affects the moisture retention capacity of the soil, but also the pH and quality of nutrients it retains. [d] A dummy variable relating to the soil quality for crop production; it equals 1 if the plot has no or only a slight constraint on field management where the constraints are provided by the harmonised world soil database and relate to soil texture, effective soil, depth/volume, and soil phases. [e] The number of household members under 12 years old divided by the number of household members 12 or more years old. [f] Equals 1 if improvement of the access in the past five years is reported at the community level and 0 otherwise.

transaction costs and no access to credit, Malawian farmers are trapped into subsistence farming (Harashima, 2008; Zeller, Diagne, & Mataya, 1997).

Agrarian commercialisation which has been promoted for over more than two decades is mainly associated with growing burley tobacco which constitutes more than 90 per cent of the country's export revenue. Among the few other products farmed in the country are cotton, sugarcane, maize and tea. Food self-sufficiency is based on planting maize. A considerable part of the Malawian population does not have a varied diet and over-relies on maize. Food insecurity and malnutrition have been a main concern and are among the most important policy issues in Malawi (Harrigan, 2003; Peters, Herrera, Braun, & Kennedy, 1994; WFP, 2012). Agricultural commercialisation and its controversial nutritional effects have been the focus of research since the 1990s. The literature points out food insecurity induced by displacing maize by tobacco for some (Harrigan, 2003) and structural constraints underlying the choice to avoid cash crop adoption for others. Even if agricultural income is higher, it is not necessarily related to improved nutritional outcomes since households may not have an opportunity to commercialise their output or buy products because of missing markets (Dorward & Kydd, 2004; Wood, Nelson, Kilic, & Murray, 2013). Food insecurity related to market failures in Malawi was stated by the literature based on the 1990s data and an institutional context characterised by numerous structural transformations (Harrigan, 2003); yet, some positive changes of the structure of input marketing and crop marketing, and an increasing role of private traders have been acknowledged since the 2000s (Chinsinga, 2011; Harrigan, 2003).

2.1. Nutrition-related outcomes

Carletto et al. (2015) review various measures used in studies exploring the link between agriculture and nutrition, point out their interrelation and suggest to use a set of measures. As mentioned above, we consider two different nutrition-related outcomes: household annual per capita food expenditures representing market food consumption and the food consumption score (FCS), representing diet quality of the households.

Food expenditures provided by the data are built using a survey based on the Laspeyres price index to adjust for difference in price levels across regions and, when possible, location implying that differences in expenditures across the population relate to differences in consumption rather than price variations.

Food expenditure analysis is motivated by its monetary nature and therefore its direct relationship to commercialisation of the household's production: commercialisation is meant to increase the household's monetary budget; food expenditure analysis is indicative therefore relative to the possibility to satisfy higher demand for food coming from displaced maize production and/or more cash available. Higher expenditures under cash cropping regime versus food only regime implies, thus, food security.

While indicative of markets availability and food security, food expenditure analysis would be insufficient for understanding the nutritional impact of cash cropping. Higher food expenditures are necessary but not sufficient to warrant better nutrition: they may allow compensation for the household's food production displacement without improving its nutrition under cash cropping; they can also enhance nutrition in terms of greater food consumption without improving quality and the variety of the household's diet. We therefore use the food consumption score to investigate the relationship between cash crop adoption and nutrition quality.

The food consumption score (FCS) is a composite score capturing intensity and dietary diversity of food consumption as well as the relative nutritional importance of different food groups. Following WFP (2008) and Jones, Shrinivas, and Bezner-Kerr (2014), the score is built using data on the household's consumption of nine food groups (staple grains and tubers, pulses, vegetables, fruits, meat and fish, dairy, oil and sugar) during the past seven days. Consequently, the FCS reflects the quality and diversity of the household's diet, which is of interest since higher incomes may lead a household not only to acquire more food but to have a diet of higher quality and of greater variety (Hoddinott, 2012).

Food expenditures and FCS are strongly positively correlated (0.4 correlation): higher food expenditures entail higher dietary diversity for some households. Yet, the correlation is not perfect since the FCS is inherent not only to the expenditures the households can spend on marketed food but also to the household's own food production.

The FCS is also positively correlated with the diversity of food expenditures: its correlation with the number of different food items purchased on the market is approximately 0.6. This shows further that a great deal of the FCS variation among Malawian households comes from the variety of food purchases while a part of this variation should relate to own production.

Initial empirical evidence of some differences in nutrition-related outcomes between cash croppers and food only growers is reported in Table 1. The descriptive statistics imply that the outcomes, on average, are higher among cash croppers than among food only growers. The gaps are weak (roughly 7% in terms of food expenditures and 2% in terms of dietary diversity) but statistically significant. Yet, unconditional average differences are weakly informative about the effects of cash cropping and not informative about the distribution of the effects across the agrarian population and about households' marketing opportunities and related decision-making regarding commercialisation. The next section provides the framework based on a model with essential heterogeneity (Heckman et al., 2006) allowing us to obtain more compelling parameters to analyse.

3. Framework and concepts

Let F_i be a nutrition-related outcome of household i. In the analysis below it is either the logarithm of food expenditures or the logarithm of the food consumption score FCS described in Section 2. The outcome F_i is a function of the household's wealth and income represented mainly by agricultural production value, the household's preferences which are conditional on household's composition and education, and market conditions (prices) faced by household i: $F_i = F_i(X_i, U_i)$ where X_i is the vector of observable arguments and U_i is unobserved heterogeneity relating to market conditions and preferences not captured by the observable data.

Households' food consumption patterns vary under different cropping regimes: cash crop adoption, and sowing only food crop. These regimes are indexed by C_i, an indicator which takes the value of 1 for the cash crop option and 0 for the only food one. Production decisions on the crop portfolio might depend on productivity under different regimes which is contingent on the household's agricultural inputs, endowments, and prices represented by observable data Z_i and the unobservable V_i. The populations of cash croppers and only food growers may have an unequal distribution of X and different consumption patterns. The elasticity/semi-elasticity of the outcomes, $\partial F/\partial X$, can therefore be different under the two different regimes: the marginal effect of X is γ_0 under $C = 0$ (no cash cropping) and γ_1 under $C = 1$ (cash cropping).

Further, incompleteness of markets inherent to rural economies and more generally developing countries entails non-separability between production and consumption decisions of the households: unlike production decisions under perfect markets resulting in agricultural profit maximisation, production decisions under market failures might depend on consumption patterns (Bardhan & Udry, 1999). Dependence between decisions on food consumption and crop portfolio is explicitly modelled by allowing for essential heterogeneity conceptualised below.

The model with essential heterogeneity derives from the generalised Roy model (Heckman et al., 2006) and can be presented in a switching format:

$$F_i = \gamma_0 X_i + (\gamma_1 - \gamma_0) \cdot C_i \cdot X_i + U_i \tag{1}$$

with

$$U_i = U_{0i} + C_i \cdot (U_{1i} - U_{0i}), \tag{2}$$

$$C = \begin{cases} 1, & if\, P_i(Z_i) - Vi \geq 0 \\ 0, & if\, P_i(Z_i) - Vi < 0 \end{cases} \tag{3}$$

and

$$U_i \not\perp V_i \tag{4}$$

The cash crop adoption decision is guided by selection rule Equation (3) where $P_i(Z_i)$ and V_i[2] represent the propensity scores of planting cash crops in terms of observables and the propensity scores of not planting cash crops in terms of unobservable determinants of cash crop adoption respectively.

Equations (2) and (4) describe the essential heterogeneity. Equation (2) implies that the unobserved part of the outcome of household *i*, U_i, is also switching between two regimes: $U_i = U_{0i}$ if household *i* chooses to abstain from cash cropping, and $U_i = U_{1i}$ if household *i* chooses to plant a cash crop. The differential $(U_{1i} - U_{0i})$ is household – specific implying that unobserved gains (or losses in case of negative differentials) from cash cropping, $(U_{1i} - U_{0i})$, are heterogeneous among the households. The differentials have interesting implications in our framework. While the unobserved heterogeneity of nutrition-related outcomes of households might relate to heterogeneous preferences for expenditures or diet and disparate prices faced by different households, the household-specific differential $(U_{1i} - U_{0i})$ does not relate to any of these. $(U_{1i} - U_{0i})$ is not affected by the prices since household *i* faces the same market prices under any cropping regime and have the same preferences and needs depending rather on its composition and education. Non-zero differentials imply, therefore, access to different consumption bundles under different regimes while holding agricultural income fixed. In the framework of missing markets, such accessibility difference indicates the possibility to spend cash revenue gained through agricultural commercialisation. A positive differential implies, therefore, the availability of markets to the household and the possibility to optimise its consumption bundle. On the contrary, a negative differential may imply worsening of nutrition under a cash cropping regime which can be due to the impossibility of selling the output and/or buying goods, or lower preferences for food as compared to other goods that may include agricultural inputs and household assets.

The second aspect of the essential heterogeneity is stated by Equation (4) and consists of correlation between the selection process into cash cropping and relative gains or losses from cash crop adoption, meaning that the households can sort into different regimes considering their gains from choosing one crop portfolio rather than another. A positive (negative) sorting on the gain related to positive (negative) correlation between the gains from cash cropping and the likelihood of cash crop adoption means a higher (lower) likelihood to plant cash crops for the households whose nutrition improves (worsens) from cash crop adoption.[3] Detecting such a correlation represents an empirical test of non-separability of production and consumption decisions and is instructive relative to the market barriers and the corresponding households' behaviour.

Specifically, a positive sorting on the gain would imply that the households accessing markets are more likely to diversify their crop portfolio by cash crops while those facing stronger barriers are less likely to adopt cash crops. Such a pattern signals first the presence of barriers to selling agricultural output and/or buying marketed food using the cash revenue; second, it means that the intensity of the barriers faced vary within the rural population; and third, that households are able to optimise their crop portfolios by anticipating the barriers faced and making their decisions on cash crop adoption taking into account eventual gains or losses.

A zero sorting on the gain (uniformly distributed returns to unobservables) would imply no relationship between the production decision process and differences in food consumption patterns under different cropping regimes signalling separability between production and food consumption decisions. Uniformly distributed positive differentials would imply uniform access to marketing and uniform benefit from agricultural commercialisation with no linkage to the production decisions while uniformly distributed negative differentials would imply a uniform loss from commercialisation and inability of the households to anticipate it.

Finally, a negative sorting on the gain would imply the strongest gains from commercialisation among those who are the most likely to abstain from it. Such a pattern can signal other kinds of barriers faced by the households that sow only food crops while having the possibility to sell the non-food agricultural outputs and/or purchase marketed food. These can be barriers related to a lack of inputs, financial constraints or risk aversion limiting households' choice relative to crop portfolios.

4. Model outcomes and their estimation

The households' differentials $(U_{1i} - U_{0i})$ are unobserved and cannot be predicted. However, their distribution in the population can be inferred from the distribution of the marginal treatment effects, *MTE*, and summarised by the aggregate differentials known in the literature as average treatment effects. The marginal treatment effect in our context is the return to cash crop adoption conditional on the household's observable and unobservable characteristics *X* and *V*:

$$MTE(X_i, V_i) = E[F_{1i} - F_{0i}|X_i, V_i]$$

It is identified by $MTE(X_i, P_i) = \frac{\partial E[F_i|X_i, P_i]}{\partial P_i}$ (Heckman et al., 2006), the marginal gap of the household's nutrition-related outcome associated with an infinitesimal change of the likelihood to adopt cash crops $(P_i = P(Z_i))$.

In the framework of Equations (1) – (4), the *MTE* splits into two parts related to observable and unobservable covariates[4]:

$$MTE(X_i, V_i) = (\gamma_1 - \gamma_0)X_i + MTE^V(V_i) \tag{5}$$

where $MTE^V(V_i) = E[U_{1i} - U_{0i}|V_i = P_i]$. The MTE^V definition implies that the MTE^V can be evaluated at the range of values where $V = P(Z)$, that is at the points conforming to the propensity scores of cash cropping predictable from Equation (3) and corresponding to the points where farmers are indifferent between the crop portfolios. Identification of the MTE^V requires instrumental variables among the covariates *Z* of selection Equation (3) in order to ensure variation of the propensity score unrelated to variation of the main outcomes. The range of propensity scores provides a set of MTE^V that can be associated with various levels of the likelihood of cash crop adoption. Tracing out the MTE^V within the support space of the propensity scores of not planting cash crops, *V*, enables detection of the different schemes conceptualised in Section 3 and testing the presence of market barriers, their heterogeneity, and the farmer's ability to anticipate nutrition-related gains or losses from cash cropping and optimise the crop portfolio accordingly.

Moreover, integrating the MTE^V over the support of the propensity scores defined by $V = P$ and weighting differently individual observations depending on their contribution to the likelihood of cash cropping allows for identification of the average treatment effects (Heckman et al., 2006) $(ATE^V = E(U_1 - U_0),\ TT^V = E(U_1 - U_0|C = 1)$ and $TNT^V = E(U_1 - U_0|C = 0))$. This is otherwise impossible given multiple heterogeneity (heterogeneity of the households across the groups and heterogeneity of the returns from cash crop adoption) and multiple selection (selection on the levels of the outcomes and the gains from cash crop adoption). The three effects representing the same gaps differ by the population referred: unconditional[5] ATE^V refers to the average differential in the whole population of the farmers, TT^V and TNT^V refer to the subpopulations of the farmers planting cash crops or refraining from it respectively. Their identification is conditional on the completeness of the common support of *P* (propensity scores predictions encountered among both, cash cropping households and households not planting any cash crop), that is the common support overlay of the unit interval $[0; 1]$.

Furthermore, the total average effects can be calculated by summing up the treatment effects related to the population's observables $(ATE^X = (\gamma_1 - \gamma_0)E(X),\ TT^X = (\gamma_1 - \gamma_0)E(X|C = 1)$ and $TNT^X = (\gamma_1 - \gamma_0)E(X|C = 0))$ and the treatment effects related to unobservables (ATE^V, TT^V, and TNT^V): $ATE = ATE^X + ATE^V$, $TT = TT^X + TT^V$ and $TNT = TNT^X + TNT^V$.

Estimation of the average effects provides additional parameters for building the analysis and testing the presence of market barriers and farmers' behavioural schemes. Specifically, $TT > ATE > TNT$ indicates positive sorting on the gain, $TT = ATE = TNT$ indicates zero sorting on the gain, and $TT < ATE < TNT$ points towards negative sorting on the gain on average in the population. The ordering of the components relating to observables and unobservables permits the analysis in terms of observable and unobservable determinants of the main outcomes respectively.

The model parameters are estimated using an estimation procedure detailed in Heckman et al. (2006). First, the propensity score, P, is predicted using a probit model associated with selection Equation (3). Next $(\gamma_1 - \gamma_0)$ and γ_0 are recovered using a double residual semiparametric regression. MTE^V are evaluated using a local linear regression at different points of the common support of P. ATE^V, TT^V, and TNT^V are evaluated by integrating the MTE^V over the full common support. ATE^X, TT^X, and TNT^X are calculated at the means of the whole sample, sample of cash croppers and sample of only food growers respectively using the estimates of $(\gamma_1 - \gamma_0)$. The total effects ATE, TT, and TNT aggregate ATE^V, TT^V, TNT^V and ATE^X, TT^X, and TNT^X. The standard errors are obtained via the bootstrap method.

5. Results

The estimated parameters of the probit regression applied to selection Equation (3) and the semiparametric regression employed to estimate Equation (1) are reported in Table 2 and discussed in Section 5.1. The key results of the analysis and the treatment effects discussed in Section 5.2, are displayed on Figures 2 and 3 and in Table 3.

5.1. Observable determinants of cash crop adoption and nutrition-related outcomes

The production decision in terms of the crop portfolio choice might depend on the productivity under different cropping regimes which is conditional on the household's agricultural inputs, endowments, market and agroecological conditions. Accordingly, the vector of explanatory variables includes acreage, controls for topographic wetness conditions and soil quality relative to crop production, and climatic zones. Observable heterogeneity is captured by demographic and human capital variables such as household composition (the size and child-dependency ratio), and the household head's age, gender, and education. Market conditions are represented by community-level variables available, such as indicators on stability of access to manufactured and staple goods and distances to population centres, borders, and tobacco auction floors; inclusion of district dummies (there are 28 districts in Malawi) allows control for regional effects including local price levels.

The results show that larger land holdings and favourable agro-climatic conditions increase the likelihood of cash crop adoption; the average marginal effect of increasing acreage is 8.5 per cent. The likelihood also varies among households of different composition. Larger and male headed households are more likely to diversify their crop portfolio by adopting cash crops (the average marginal effects are 8 and 0.6% respectively) while a higher number of children per adults disfavours cash cropping.

Positive dynamics of access to staple foods and lower distance to a tobacco auction floor encourages cash cropping. Access to staple food plays a considerable role with an estimated marginal effect of 6 per cent. Stable access to staple foods is likely to increase food security and consequently willingness to sow cash crops. Stable access to manufactured goods is not found to stimulate cash crop adoption and commercialisation; food security is more important for farmers than the possibility of purchasing other market goods. The marginal effect of the distance to tobacco auction floors is also significant: the farther away the household is from a tobacco auction floor, the less likely it will engage in production of non-food crops. Specifically, at the average level of distance (about 77 km), a one standard deviation decrease implies a 7 per cent increase of the propensity score to plant non-food items (from 17 to 24%).

The probit estimates are employed to predict the propensity scores P of adopting cash crops. The scores are used at the semiparametric estimation of the nutritional outcomes equation and nonparametric estimation of the treatment effects associating them with V, the propensity score of refraining from cash cropping. Identification of the population average effects ATE^V, TT^V and TNT^V, whose estimates are discussed in the next subsection, is conditioned by the requirement of fullness of the common support.[6] Figure 1 shows that the common support of the score distributions among cash croppers and food only growers is the whole unit interval $[0; 1]$ meaning the full support and validating the assumption.

Table 2. Estimates of the selection and nutrition-related equations

	Probit regression estimates		Semiparametric regression estimates			
	Cash crop adoption		Food expenditures		FCS [a]	
Covariates	Base estimates	Average marginal effects	$\hat{\gamma}_0$	$(\hat{\gamma}_1 - \hat{\gamma}_0)$	$\hat{\gamma}_0$	$(\hat{\gamma}_1 - \hat{\gamma}_0)$
Logarithm of harvest value [b]			.063***	.096**	.025***	.032**
			(0.01)	(0.04)	(0.00)	(0.01)
Household agro-ecological conditions						
Acreage, Ha	.467***	.085***				
	(0.06)	(0.00)				
Potential wetness index (PWI) [c]	.147***	.010***				
	(0.04)	(0.00)				
PWI squared	−.003***					
	(0.00)					
Workability constraint index [d]	.120*	.022*				
	(0.07)	(0.01)				
Indicators of agro – ecological zones	Yes	Yes				
Demographic characteristics						
Head's age	−.009***	−.002***	.001	−.003	−.001***	−.002
	(.001)	(.000)	(.001)	(.003)	(.000)	(.003)
Female headed household	−.439***	−.080***	−.004	−.026	−.031***	.172
	(.055)	(.009)	(.029)	(.193)	(.010)	(.162)
Household size	.036***	.006***	−.120***	.100***	.003*	−.013
	(.009)	(.002)	(.005)	(.019)	(.002)	(.069)
Child dependency ratio [e]	−.060**	−.011***	−.078***	−.209***	−.006	−.114
	(.030)	(.006)	(.014)	(.076)	(.006)	(.076)
Head's education						
None	–	–				
Primary	.119*	.022**	.125***	−0.135	−.126***	−0.113
	(.062)	(.010)	(0.033)	(0.120)	(0.040)	(0.113)
Secondary or Tertiary	−.165**	−.030***	.344***	−.207	−-.099**	.037
	(.065)	(.010)	(.030)	(.156)	(.049)	(.132)
Community features						
Improved manufacture [f]	−.028	−.005				
	(.095)	(.009)				
Stable staple [f]	.319***	.058***				
	(.093)	(.011)				
Distance to population center, Km	.003	.000*	.000	−.002	.001	.001
	(.002)	(.000)	(.001)	(.003)	(.001)	(.002)
Distance to border, Km	.009***	.002***	−.004***	.006***	.001	.003
	(.002)	(.000)	(.001)	(.001)	(.001)	(.001)
Distance to tobacco auction floor,	−.007***	−.001***				
Km	(.002)	(.000)				
District dummies [g]	*Yes*		*Yes*		*Yes*	
Constant	−1.097***					
	(.311)					
Observations	9,514	–	9,441		9,173	
(Pseudo) R2	0.27	–	0.28		0.24	

Notes: Standard errors clustered at the village level, *** (1%), ** (5%), * (10%). [a–f] see notes for Table 1. [g] Malawi is divided into 28 districts.

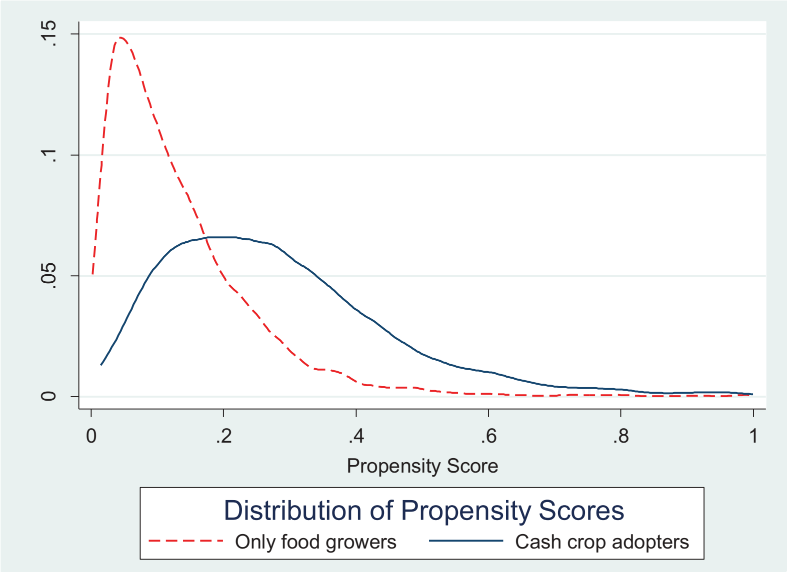

Figure 1. Propensity scores of cash crop adoption.

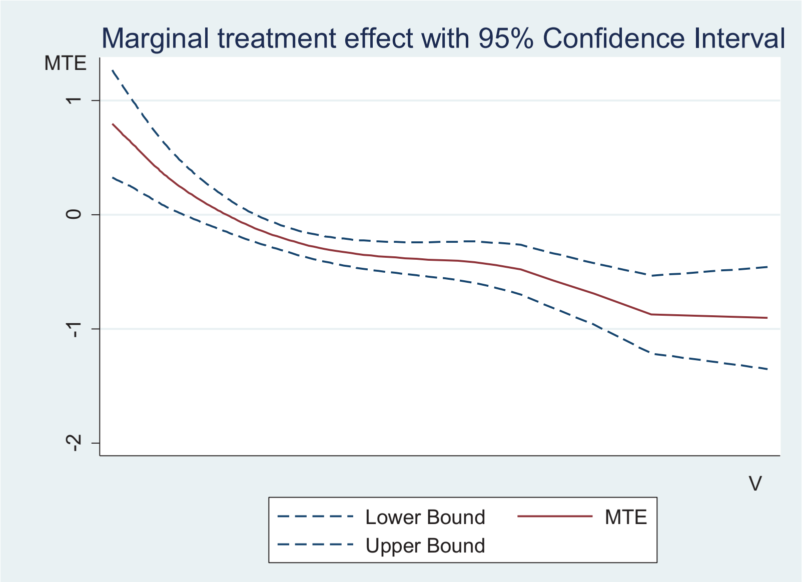

Figure 2. MTE^V, food expenditures.

The set of covariates of the nutrition-related outcomes (logarithms of food expenditures and food score) overlap with the set of probit regressors. Additionally, it includes the harvest value. On the other hand, the productivity factors and distance to tobacco auction floor are excluded from the semiparametric regression since they might affect nutrition-related outcomes only through the harvest value but are not direct determinants of food demand and expenditures. Indicators of positive dynamics of access to manufactured goods and staple food are also excluded since at the proximity to subsistence thresholds, as is the case in Malawi, food expenditures are the short-run solution responding to

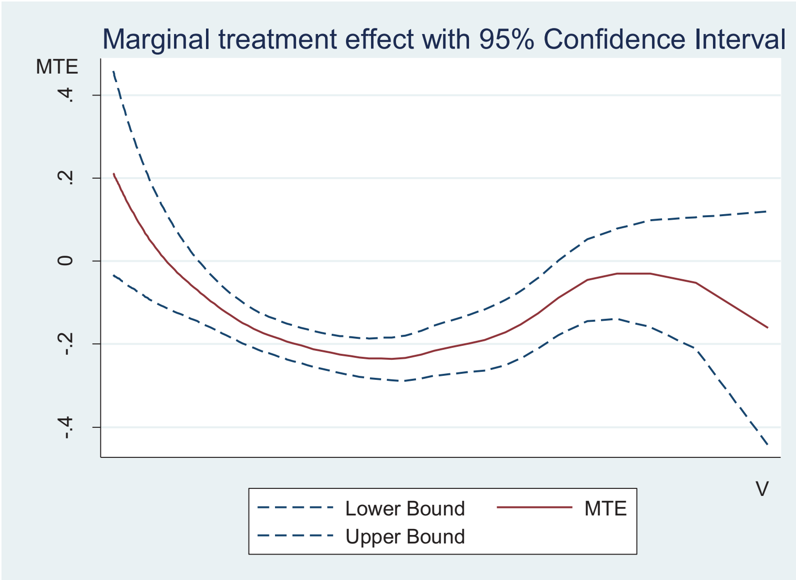

Figure 3. MTE^V, dietary diversification (FCS).

household necessities rather than a solution coming out of long-term utility optimisation. Exclusion variables allow for identification of the marginal treatments effects – MTE^V reported in the next subsection – by inducing variations of the propensity scores of cash cropping unrelated to food consumption determinants.

The coefficients corresponding to households that abstain from cash cropping are presented in columns 3 (food expenditures estimates) and 5 (FCS estimates) labelled by γ_0 in Table 2. The differences in coefficients between households that do not plant cash crops (γ_0) and those who do (γ_1) are reported in columns 4 (food expenditures estimates) and 6 (FCS estimates) labelled by $(\gamma_1 - \gamma_0)$.

The harvest elasticity of food expenditures and nutrition diversification represented by FCS is weak (0.06 and .025 respectively) but highly statistically significant. Both semiparametric regressions show more than double harvest elasticity among the cash cropping households (the corresponding differences are 0.096 and 0.032 respectively). According to Table 1, cash cropping households operate at a considerably higher level of production output. Higher income elasticity among them implies that at a higher level of production output, the households can afford to spend a higher income share on food and diversify their nutrition more intensely. Empirical evidence of positive income impacts on the per capita food expenditures also comes out through positive association of food expenditures with higher education of the household's head and smaller number of dependents reflected by negative impacts of larger size of the household and number of children per adults.

The estimates associated with harvest values might be underestimated in the case of positive correlation between production output and the error term of the food consumption equation, U_i. The endogeneity of the production output may be due to reverse causality between nutritional and agricultural outcomes or common unobservable factors impacting both outcomes: at low levels of income, higher production output might induce better nutrition while better nutrition might lead to higher productivity (Deolalikar, 1988; Strauss, 1986). Positive correlation between food outcome unobservables and production outcomes could bias downward the harvest income elasticity of nutrition-related outcomes. However, eventual endogeneity does not yield a threat to validity of the key parameters relating to various treatment effects discussed in Section 5.2: independence between exclusion variables and (U_1, U_0) is conditional on X implying that identification is not impacted by eventual correlation between (U_1, U_0) and X. Endogeneity of the covariates of the nutritional equations is therefore harmless. This is less restrictive as compared to conventional parametric approaches requiring that not only the exclusion variables but also explanatory variables X be independent of (U_1, U_0) (Heckman & Vytlacil, 2005).

5.2. Food consumption differentials and crop portfolio decisions

The central results of the analysis, the estimates of the average treatment effects which are presented in Section 4 and reported on Figures 2 and 3 and in Table 3, are instructive in two ways.

First, while unconditional differences in nutritional outcomes of cash croppers and food-only growers reported in Table 1 are weak (roughly 7% in terms of food expenditures and 2% in terms of dietary diversity), conditional differentials reveal non-zero nutritional and food expenditures effects of agricultural commercialisation. More interestingly, the effects are heterogeneous across the population as reflected by the inequality of *TT*, *ATE* and *TNT* (Table 3): cash cropping benefits or hurts disparately different tiers of the agrarian population.

Specifically in terms of food expenditures, the total population differentials range from − 11 per cent for the population refraining from cash cropping as reported by *TNT* to 45 per cent for the population of cash croppers as reported by *TT*; the resulting total population average effect, *ATE*, is about 15 per cent. A positive (negative) differential implies a positive (negative) average effect of cash cropping. In terms of dietary diversification, the average population effect mirrors the descriptive statistics coming close to zero; however, the average differentials corresponding to the subpopulations of households abstaining from cash cropping, *TNT*, and cash crop adopters, *TT*, imply a 9.5 per cent average nutritional loss from commercialisation in the first subpopulation and 8 per cent average gain in the second one. The decomposition of the average effects into the components relating to observable and unobservable determinants of the outcomes show that positive gains for the cash croppers relate to a large extent to observables (food expenditures $TT^X = 26\%$, FCS $TT^X = 7.9\%$). This is mainly due to higher harvest value elasticity of the outcomes coming along with higher harvest mean value among cash crop adopters. The results suggest a transmission from higher agricultural incomes to higher nutrition-related expenditures and a more varied diet.

To a large extent, the similarity of the aggregate results relating to food expenditures and FCS is due to interrelationship between the two variables. As discussed in Section 2, food expenditures and the FCS are strongly but not perfectly correlated. Displacement of food production by cash crops might reduce dietary diversity for those who crop various staple food items. Cash crop adoption might, therefore, have ambiguous effects on the FCS: increasing income from commercialisation might favour more varied consumption of market food but less varied consumption of self-produced food. The below analysis of differentials and essential heterogeneity in terms of unobservables reveals considerable differences between the commercialisation effects on food expenditures and dietary quality.

Table 3. Differentials of nutritional outcomes

	Total	*Observables*	*Unobservables*	$SB_{1\to 0}$ [b]	$SB_{0\to 1}$ [c]
Food expenditures				−.226	−.164
TT	.445***	.264***	.181*		
ATE	.155**	.127**	.028		
TNT	−.109***	.099***	−.208***		
FCS [a]				−.038	−.82
TT	.082*	.079*	.003		
ATE	−0.004	0.032	−.036		
TNT	−.095***	.022***	−.117***		

Notes:Population effects are estimated using sample weights.[a] See notes for Table 1.[b] Selection term showing the difference in outcome between cash croppers and only food growers should they abstain from cash cropping. [c] Selection term showing the difference in outcome between only food growers and cash croppers should they plant a cash crop.

5.2.1. Essential heterogeneity of food expenditures. In terms of unobservables, both households adopting cash crops and refraining from it have on average roughly 20 per cent higher food expenditures under the chosen portfolio than under the counterfactual one (a negative *TNT* implies a gain from refraining from cash cropping). The unobservable effects are efficiently illustrated by Figure 2 tracing out the marginal treatment effects, MTE^V, versus the propensity to refrain from cash cropping, V. As mentioned in Section 4, TT^V and TNT^V are obtained by integrating MTE^V over the range of the propensity V while weighting individual observations differently: $TT^V(TNT^V)$ attributes stronger weights to observations with higher (lower) likelihood of cash cropping. As discussed in Section 3, given the control for harvest value capturing income effect on food consumption, non-zero differentials TT^V and TNT^V of food expenditures imply that under different cropping regimes the same households access different consumption bundles due to different marketing possibilities.[7] The positive TT^V implies availability of functioning markets for cash crop adopters and the possibility to optimise their consumption bundles and satisfy higher demand for food coming from displaced maize production and/or larger budget availability. The negative TNT^V signals a reduction of food consumption under the cash cropping regime for the food only growers which can be due to market barriers relative to selling the output and/or purchasing marketed goods.

Figure 2 conveys further lessons of the analysis by showing a positive sorting on the food expenditures implied by the MTE^V downward sloping curve. The figure shows that heterogeneous differentials are correlated with cash-crop adoption likelihood: households which are most likely to plant cash crops based on their unobservable characteristics (at low propensity V) have higher levels of food expenditures and therefore better marketing opportunities. The differentials for those who are less likely to adopt cash cropping are negative, implying higher food consumption benefits from planting only food items. The estimates suggest that these households would have lower food expenditures should they choose to plant cash crops rather than only food items and imply therefore that they face stronger market barriers to selling agricultural outputs and/or buying marketed food using the cash revenue. In line with the downward sloping curve, the aggregate treatment effects are ordered as $TT^V > ATE^V > TNT^V$.

The results concur with the non-separability property: when making decisions relative to crop portfolios, households consider food consumption patterns. In addition, the positive sorting on the gain signals rational anticipations and decision process of agrarian households relative to their crop portfolio choice. Households are less likely to engage in cash cropping when expecting stronger market barriers and consequent food insecurity; their decisions on cash crop adoption are thus guided by food security considerations and gains in terms of market food consumption.

5.2.2. Essential heterogeneity of food consumption scores. Food consumption score analysis does not follow exactly the same pattern as above, but also shows some positive sorting on the gain: $ATE^V > TNT^V$ along with the MTE^V curve downward at low and medium values of V while showing an unclear pattern at its high values (Figure 3). Specifically, the FCS analysis implies that households keeping a food only crop portfolio would have a less varied diet (TNT^V is about -12%) should they choose to adopt cash crops. The results suggest that these households would have lower food expenditures and a less varied diet should they choose to plant cash crops rather than only food items. When refraining from agricultural commercialisation, food only growers are likely to be driven not only by market failures and food insecurity but also by anticipated deterioration of their diet's quality.

Cash cropping households, on average, would not see significant difference between two regimes in terms of dietary diversity ($TT^V \simeq ATE^V \simeq 0$) should they change their crop portfolios. Their decisions to adopt cash crops are therefore driven by an increase of marketed food consumption as reflected by food expenditure analysis; these gains in terms of food consumption are enough to compensate for a possible diet degradation of cash coppers following commercialisation while the gains do not bring unexpected improvement of the cash croppers' dietary diversity.

All in all, the results indicate non-separability between households' decisions on consumption and production. The results imply that food insecurity is a primary motive holding households back from cash crop adoption and to a traditional pattern of planting food-only crops.

Finally, the negative selection bias terms imply unobserved heterogeneity between households from different groups: each type of agrarian household would have lower levels of food expenditures and consumption scores compared to the other type should all the households choose the same crop portfolio, diversified by the adoption of cash crops or by planting only food crops.

6. Conclusions

Using Malawian data, the analysis sheds light on the inconclusiveness of previous research relative to food security and nutritional effects of agricultural commercialisation in an agrarian economy: the results obtained by applying a model with essential heterogeneity show that the effects are heterogeneous across the agrarian population due to different degrees of market barriers faced by households relative to selling their harvest and/or buying marketed goods and, consequently, various degrees of food insecurity. Adoption of cash crops benefits or hurts different tiers of the agrarian population disparately.

Furthermore, the results suggest a positive sorting on food expenditures signalling rational anticipation and a decision process of farming households relative to the crop portfolio choice. Households are also less likely to engage in cash cropping when expecting food insecurity and malnutrition; those having weaker marketing barriers are more likely to adopt cash crops. These results concur with the non-separability property: when making production decisions relative to crop portfolios, households consider food consumption patterns. More generally, the approach applied can be instructive for testing separability between production and consumption decisions of farming households.

Additionally, the results suggest a transmission from higher agricultural incomes obtained from cash cropping to higher nutrition-related expenditures and a more varied diet while the underlying harvest value elasticity is relatively weak, implying that increased incomes alone may not be sufficient for improved nutritional outcomes.

Acknowledgements

The data and code are available on request.

Funding

Computing resources used for this work are provided by the American University High Performance Computing System, which is funded in part by the National Science Foundation (BCS-1039497) see www.american.edu/cas/hpc for information on the system and its uses.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes

1. Less than 1 per cent of the Malawian households plant only cash crops.
2. By construction, V_i follows a uniform distribution over [0;1].
3. Under controlled/standardised income and price variations across the population.
4. See Heckman et al. (2006) for derivation.
5. Unconditional on C.
6. The full support requirement is the strongest in regard to the ATE^V estimation: $P(Z)$ values arbitrary close to 0 and 1 are needed; the estimation of TT^V and TNT^V is less demanding, requiring only some positive $P(Z)$ and the values close to 0 and 1 respectively (Heckman & Vytlacil, 2000).
7. Note that as specified in Section 2, food expenditures are adjusted for differences in price levels across regions and locations.

References

Bardhan, P., & Udry, C. (1999). *Development microeconomics*. Oxford: OUP.

Carletto, C., Corral, P., & Guelfi, A. (2016). Agricultural commercialization and nutrition revisited: Empirical evidence from three African countries. *Food Policy*. doi:10.1016/j.foodpol.2016.09.020

Carletto, G., Ruel, M. T., Winters, P., & Zezza, A. (2015). Farm-level pathways to improved nutritional status: Introduction to the special issue. *The Journal of Development Studies*, *51*, 945–957. doi:10.1080/00220388.2015.1018908

Chinsinga, B. (2011). *Agro- dealers*, subsidies and rural market development in Malawi: A political economy enquiry. (Future Agricultures Working Paper *No*. 31, Institute of Development Studies, University of Sussex).

Deolalikar, A. B. (1988). Nutrition and labor producvivity in agriculture: Estimates for rural South India. *The Review of Economics and Statistics*, *70*, 406–413. doi:10.2307/1926778

Dillon, A., McGee, K., & Oseni, G. (2015). Agricultural production, dietary diversity and climate variability. *The Journal of Development Studies*, *51*, 976–995. doi:10.1080/00220388.2015.1018902

Dorward, A., & Kydd, J. (2004). The Malawi 2002 food crisis: The rural development challenge. *The Journal of Modern African Studies*, *42*, 343–361. doi:10.1017/S0022278X04000229

Goldstein, M., Mchich, K., Montalvao, J., Munoz, M., O'Sullivan, M., & Zikronah, B. (2013). *"You can't eat cotton": Evidence on crop portfolios and gender from Benin*. (Roadmap for Promoting Women's Economic Empowerment Report, United Nations Foundation).

Harashima, A. (2008). *The impact of tobacco production liberalization on smallholders in Malawi* (IDE Discussion Papers *No*. 170, Institute of Developing Economies, Japan).

Harrigan, J. (2003). U-turns and full circles: Two decades of agricultural reform in Malawi 1981-2000. *World Development*, *31*, 847–863. doi:10.1016/S0305-750X(03)00019-6

Heckman, J., & Vytlacil, E. J. (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, *73*, 669–738. doi:10.1111/ecta.2005.73.issue-3

Heckman, J. J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics*, *88*, 389–432. doi:10.1162/rest.88.3.389

Heckman, J. J., & Vytlacil, E. J. (2000). *Local instrumental variables* (NBER Technical Working Papers No.0252, National Bureau of Economic Research, Inc.).

Hoddinott, J. (2012). *Agriculture, health, and nutrition: Toward conceptualizing the linkages*. 2020 Conference Book: Leveraging agriculture for improving nutrition and health. Washington, DC.:International Food Policy Research Institute

Jones, A. D., Shrinivas, A., & Bezner-Kerr, R. (2014). Farm production diversity is associated with greater household dietary diversity in Malawi: Findings from nationally representative data. *Food Policy*, *46*, 1–12. doi:10.1016/j.foodpol.2014.02.001

Kumar, N., Harris, J., & Rawat, R. (2015). If they grow it, will they eat and grow? Evidence from Zambia on agricultural diversity and child undernutrition. *The Journal of Development Studies*, *51*, 1060–1077. doi:10.1080/00220388.2015.1018901

Masanjala, W. H. (2006). Cash crop liberalization and poverty alleviation in Africa: Evidence from Malawi. *Agricultural Economics*, *35*, 231–240. doi:10.1111/agec.2006.35.issue-2

Peters, P., Herrera, M., Braun, J., & Kennedy, E. (1994). *Tobacco cultivation, food production, and nutrition among smallholders in Malawi*. (Agricultural commercialization, economic development, and nutrition, 309–327). Baltimore, MD: Johns Hopkins University Press.

Slavchevska, V. (2015). Agricultural production and the nutritional status of family members in Tanzania. *The Journal of Development Studies*, *51*, 1016–1033. doi:10.1080/00220388.2015.1018906

Strauss, J. (1986). Does better nutrition raise farm productivity? *Journal of Political Economy*, *94*, 297–320. doi:10.1086/261375

Strauss, J., & Thomas, D. (1995). Human resources: Empirical modelling of household and family decisions. In J. Behrman & T. N. Srinivasan (Eds.), *Handbook of development economics* (Vol. *3A*, pp. 1883–2023). Amsterdam: Elsevier.

Von Braun, J., & Kennedy, E. (1994). *Agricultural commercialization, economic development, and nutrition*. Baltimore, MD: Johns Hopkins University Press.

WFP. (2008). *Food consumption analysis: Calculation and use of the food consumption score in food security analysis*. United Nations World Food Programme.

WFP. (2012). *Malawi - Comprehensive food security vulnerability analysis and nutrition assessment*. World Food Programme.

Wood, B., Nelson, C. H., Kilic, T., & Murray, S. (2013). *Up in smoke? Agricultural commercialization, rising food prices and stunting in Malawi*. (World Bank Policy Research Working Paper No.6650). Washington, DC: World Bank.

Zeller, M., Diagne, A., & Mataya, C. (1997). *Market access by smallholder farmers in Malawi: Implications for technology adoption, agricultural productivity, and crop income* (FCND Discussion Papers No.35). Washington, DC: International Food Policy Research Institute.