Household Food Demand in Nigeria: an Application of Multivariate Double-hurdle Model

D. Akerele*, R. Tiffin, & C. S. Srinivasan

8th April 2013

Paper prepared for presentation at the 87th Annual Conference of the Agricultural Economics Society, University of Warwick, United Kingdom

8 - 10 April 2013

Corresponding author: D. Akerele (e-mial: d.akerele@pgr.reading.ac.uk)

Household Food Demand in Nigeria: an Application of Multivariate Double-hurdle Model

Abstract

A Bayesian approach was used to analyse household demand for staples in Nigeria within the framework of a multivariate double-hurdle model to account for censoring emanating from non-participation decisions. We demonstrate how to impose identification constraints in the probit equations and introduce a straightforward way of mapping observed and latent shares in the demand (share) equations to satisfy adding up restrictions. Demands for cereals, beans and tubers are own-price inelastic with values close to unity in the lowest income quintile. Cross-price elasticities indicate demand patterns characterised by a mix of gross substitutability and complementarity relationships among staple food subgroups. Presence of children and adolescents in the household as well as rural-urban and zonal (regional) differences exact significant influence on demand for staple foods. Total expenditure elasticities on staple food subgroups decline with higher income levels in line with Engel's law. Cereals and tubers are all necessary goods in the middle and highest income quintiles. However, they are luxury goods in the lowest income quintile. Our findings suggest that economic growth coupled with targeted interventions such as cash or food stamp transfer programmes are crucial for improved consumption of major staples and nutrition among households.

Key words: Bayesian approach, food demand, double-hurdle model, staples, Nigeria

JEL Code: D12

1 Introduction

Despite substantial efforts to increase food supply through domestic production and massive importations over the past years, the rates of food insecurity and malnutrition in Nigeria is still alarming. Recent studies have reported food insecurity prevalence between 60 and 79 per cent among households (Arene and Anyaeji, 2010; Orewa and Iyangbe, 2010; Olayemi, 2012; Omuemu et al., 2012) with incidence of malnutrition and related disorders spanning from 26.67 to 84.30 per cent in different parts of the country (Akinyele, 2009; Goon et al., 2011; Aliyu et al., 2012; Ubesie et al., 2012).

Previous efforts to enhance food security achieved limited successes as most of the interventions concentrated more on the supply side of the problem with little detailed evaluation of the demand side. Soaring prices of food commodities, inadequate purchasing powers, and income inequalities have been identified, among others, as critical demand side factors stimulating food insecurity and malnutrition among the majority of households in Nigeria and other developing countries (Obayelu, 2010; Olagunju et al., 2012; FAO, 2012). The greater burdens of food insecurity and malnutrition are often endured by poor households. Over 70 percent of Nigerian households are poor and expend between 60 to 80 percent or more of their earnings on food (Fregene and Bolorunduro, 2009; Adejobi and Babatunde, 2010; Obayelu, 2010). Available statistics (national average) indicate that staples account for about 55 percent of the food budget in Nigeria (NBS, 2012) with most poor households devoting more than 60 percent of their food spending to staples (Ashagidigbi et al., 2012; Ogunniyi et al., 2012).

The focus of this paper is on household demand for staples. Staples are the main dietary sources of calories and proteins among Nigerian households. They also constitute important sources of micro-nutrients in the diets of many households (Gegios et al., 2010; Musa et al., 2012). The foregoing underscores why detailed analysis of the structure of household demand for staples is vital for food security and nutrition interventions in the country, especially if policies are to impact on households through the marketplace. And, in a highly stratified socioeconomic setting such as Nigeria, household food consumption responses to such interventions could vary by income classes. Consequently, the first objective of this study is to examine household demand for staples in Nigeria by income classes. We employ the household data from the Nigeria Living Standard Survey (2003/2004) for analysis.

One of the prominent attributes of cross-sectional (household) consumption studies is the preponderance of zero expenditure records (censoring in the response/dependent variables); and how to appropriately handle them has always been a major challenge confronting applied econometricians. A number of econometric models have been employed to handle the observed zeros with each model underpinned by different hypotheses. These include the Tobit model proposed by Tobin (1958) in which the observed zero expenditures are attributed purely to economic constraints. The model was extended to a system of equations by Amemiya (1974) and have since been applied to several consumption studies (Perali and Chavas, 2000; Dong et al., 2004; Kasteridis and Yen, 2012). Following several criticisms of the traditional Tobit model, Cragg (1971) introduced a two-step (double-hurdle) model which takes into account the fact that the observed zeros might also be linked to non-participation (abstention) decisions other than pure economic reasons. The model has also been employed in its variants (Yen and Jensen, 1996; Newman et al., 2001; Obayelu, Okoruwa and Oni, 2009; Akinbode and Dipeolu, 2012). The purchase infrequency model is another econometric model developed to account for the fact that the zero-valued expenditures might arise as a result of truncated sampling period as consumers might be consuming from stock during the survey period or would have reported positive expenditures if the survey had spanned a longer period. The model has also been applied by several workers (Kimhi, 1999; Newman et al., 2001; Tiffin and Arnoult, 2010) on consumption studies. The p-tobit model (Deaton and Irish, 1984) and the double-hurdle model referred to as the pi-tobit model (Maki and Nishiyama, 1996; Maki and Garner, 2004) have also been introduced to handle zero observations associated with genuine non-consumption, purchase infrequency and misreporting.

In most cases, survey data do not provide detailed information regarding the sources of the observed zeros. In such situations, researchers are often "compelled" to adopt a particular model on the basis of the assumptions made about the potential sources of the zero observations (Kimhi, 1999; Blisard and Blaylock, 1993)[1] In this study, we attribute the observed zeros (censoring issues) in the household food consumption data to non-participation decisions. Hence, we employ the double-hurdle model for data analysis. The second objective of the study is to contribute to the literature on censoring issues within demand systems by employing Bayesian approach to the estimation of multivariate double-hurdle (sample selection) model.

The remainder of the paper is scheduled as follows. We present the double-hurdle model in a multivariate framework in section 2 while estima-

[1]Some researchers also rely on the outcomes of statistical tests performed on a range of models deemed consistent with the data (Angulo et al., 2001; Newman et al., 2001).

tion of the model is discussed in section 3. Data for the study is described in section 4 while the results of the data analysis are discussed in section 5. The conclusions are presented in section 6.

2 Double-hurdle Model in a Demand System

We estimate demand subsystems for six food groups and four staple subgroups within a multivariate double-hurdle model (MDHM). We begin with the specification of the MDHM by defining the binary process determining participation in the market of the food groups/subgroups through a set of probit equations. Leaving out one equation as it is the practice while estimating a demand subsystem, the probit equations can be presented in the latent form as:

$$\mathbf{d}^{*} = \mathbf{X_2}\boldsymbol{\varphi} + \mathbf{v} \tag{1}$$

$$\mathbf{X_2} = I_F \otimes \mathbf{x_2} \tag{2}$$

$$\mathbf{x_2} = (x_{21}, x_{22}, ..., x_{2Q})' \tag{3}$$

where $\mathbf{d}^{*}$ is a vector of latent (underlying) variables that determine household participation; and is founded on the binary process stipulated in equation 4:

$$d_{fq} = \begin{cases} 1 & if\, d^{*}_{fq} > 0 \\ 0 & if\, d^{*}_{fq} \leq 0 \end{cases} \tag{4}$$

d_{fq} is a binary outcome equalling one if household f participates in the the market of food group/subgroup q and zero otherwise. In our analysis, we presume that participation is a purely random process such that all households have equal probability of participating in the market of a particular food group/subgroup[2]. Consequently, $\mathbf{x_2}$ is set as a vector of ones. $\boldsymbol{\varphi}$ is a vector of constants and $\mathbf{v}$ represents a vector of disturbance terms.

The consumption component of the multivariate double-hurdle model is an almost ideal demand system (AIDS) model. Similarly, dropping one of the equations in the demand subsystem, the AIDS model can be specified in the latent shares form as:

[2]We acknowledge that the assumption made might be too restrictive in characterising household's participation decisions. We had difficulties in incorporating demographic factors into the probit equations in our estimation algorithm as the number of parameters to be estimated increased "exponentially" and estimation became complicated.

$$\boldsymbol{w}^* = \mathbf{X_1}\boldsymbol{\Upsilon} + \mathbf{u} \tag{5}$$

$$\mathbf{X_1} = \mathbf{I_F} \otimes \mathbf{x_1} \tag{6}$$

$$\mathbf{x_1} = (x_{12}, x_{12}, ..., x_{1Q})^{'} \tag{7}$$

$$\mathbf{x}_{1q} = \left(1,\, lnp_{1,q},\, ...,\, lnp_{f+1,q},\, ln\left(\frac{E_q}{P_q}\right),\, S_q^{'}\right)^{'}, \tag{8}$$

$$\boldsymbol{w}^* = \left(w_{1,1}^*,\, w_{1,2}^* ...,\, w_{1,Q,}^*,\, w_{f,1}^*,\, w_{f,2}^* ...,\, w_{f,Q,}^*\right)^{'}, \tag{9}$$

$$\boldsymbol{\Upsilon} = \left(\alpha_1,\, \psi_{1,1},\, \psi_{1,2},\, ...,\, \psi_{1,m+1},\, \omega_1, \varsigma_1^{'},\, \alpha_f,\, \psi_{f,1},\, \psi_{f,2},\, ...,\, \psi_{f,m+1},\, \omega_f, \varsigma_f^{'}\right)^{'}, \tag{10}$$

$$\mathbf{u} = (u_{1,1},\, u_{1,1},\, ...,\, u_{1,Q},\, u_{f,1},\, u_{f,2},\, ...,\, u_{f,Q})^{'} \tag{11}$$

Denoted by p_{fq} is the price (index) of food group/subgroup f associated with household q, E_q is the total expenditure in the demand subsystem. Denoted by $P_q = \prod_f p_{fq}^{w_{fq}}$ is the Stone's price index while S_q represents a vector of demographic and other community/regional characteristics pertaining to household q. $\boldsymbol{\Upsilon}$ is a vector of coefficients, $\mathbf{u}$ is a vector of disturbance terms while $\boldsymbol{w}^*$ is a vector of latent shares of food groups/subgroups in the demand subsystem.

Drawing from Cragg (1971), the binary process stipulated in equation 4 provides a linkage between the observed and the latent consumption (shares) of a multivariate double-hurdle model as follows:

$$w_{fq} = w_{fq}^* \qquad if\, d_{fq} = 1\, and\, w_{fq}^* > 0 \tag{12}$$

$$w_{fq} = 0 \quad if\, d_{fq} = 0\, and\, w_{fq}^* > 0 \tag{13}$$

$$or\, if\, d_{fq} = 1\, and\, w_{fq}^* \leq 0 \tag{14}$$

$$or\, if\, d_{fq} = 0\, and\, w_{fq}^* \leq 0 \tag{15}$$

where w_{fq} is the observed share on food group/subgroup j by household q. w^*_{fq} is the associated latent share. In our studies, we assume that households consume the groups/subgroups of food commodities being examined at one time or the other. So, equations 14 and 15 are not considered.

Where positive consumption is recorded (in equation 12), it is possible to formularise the mapping structure as:

$$w_{fq} = \frac{w^*_{fq}}{\sum_{f \in C} w^*_{fq}} \qquad \forall f \in C \tag{16}$$

One key feature of equation 16 is that the latent shares of (positive consumption) adds up to unity. However, when the latent share(s) associated with zero regime(s) (see equation 13) is/are added to the latent shares of positive observation(s), the adding up property of demand is violated. The adding up (coherency) problem is resolved following Tiffin and Arnoult (2010) such that:

$$w^*_{fq} = w_{fq}\left(1 - \sum_{f \notin C} w^*_{fq}\right) \qquad \forall f \in C \tag{17}$$

where w^*_{fq} is the re-specified latent shares which satisfies the adding up restriction and $\sum_{f \notin C} w^*_{fq}$ denotes the sum of the latent share(s) associated with the observed zeros.

The following conditions are imposed on the share equations in order to conform with the properties of demand model. These include:

homogeneity conditions which requires that

$$\sum_k \psi_{fg} = 0 \qquad for\, all\, g, \tag{18}$$

symmetry conditions which requires that

$$\psi_{fg} = \psi_{gf} \qquad for\, all\, f, g \tag{19}$$

and concavity which requires that the Slutsky matrix ($\mathbf{S}$) with elements:

$$S_{fg} = \psi_{fg} + \omega_f \omega_g ln\left(\frac{E}{P}\right) - w_f \delta_{fg} + w_f w_g \tag{20}$$

is negative semi-definite. Where

$$\delta_{fg} = 1 \qquad for\, f = g\;\; and\, \delta_{fg} = 0\;\, for\, f \neq g \tag{21}$$

We follow Tiffin and Arnoult (2010) to impose homogeneity and symmetry restrictions as

$$\mathbf{R}\boldsymbol{\Upsilon}^{*} = 0 \tag{22}$$

where $\mathbf{R}$ is an r x F (F+2) matrix specifying the restrictions and $\boldsymbol{\Upsilon}^{*}$ denotes the constrained version of matrix $\boldsymbol{\Upsilon}$. Incorporating the restrictions in our model requires a re-parametrisation of the model. First, by specifying a *(kF-r)* x *kF* orthonormal matrix as:

$$\mathbf{R}\mathbf{R}_{\perp}^{'} = 0 \tag{23}$$

$$\mathbf{R}_{\perp}\mathbf{R}_{\perp}^{'} = \mathbf{I} \tag{24}$$

The restricted version of ($\boldsymbol{\Upsilon}$) can be stated as:

$$\boldsymbol{\Upsilon}^{*} = \mathbf{R}_{\perp}^{'}\check{\boldsymbol{\Upsilon}} \tag{25}$$

where $\check{\boldsymbol{\Upsilon}}$ is a $(kF - r)$ x 1 vector of distinct parameters. Substitution of equation 25 into equation 5 results into its restricted version specified as:

$$\mathbf{w}^{*} = \mathbf{X_1}\mathbf{R}_{\perp}^{'}\check{\boldsymbol{\Upsilon}} + \mathbf{u} \tag{26}$$

$$\mathbf{w}^{*} = \mathbf{A}\check{\boldsymbol{\Upsilon}} + \mathbf{u} \tag{27}$$

where

$$\mathbf{A} = \mathbf{X_1}\mathbf{R}_{\perp}^{'} \tag{28}$$

The AIDS model is estimated based on equation 27 but the restricted parameters are regained from equation 25.

3 Bayesian Estimation

Prior to estimation, the sets of probit and demand equations are treated in the form of seemingly unrelated regressions (SUR) as follows:

$$\boldsymbol{y}^{*} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{e}, \tag{29}$$

where

$$\boldsymbol{y}^{*} = \left(\boldsymbol{w}^{*'}, \mathbf{d}^{*'}\right)^{'}, \quad \boldsymbol{X} = \begin{pmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{X_2} \end{pmatrix}, \quad \boldsymbol{\beta} = \left(\check{\boldsymbol{\Upsilon}}', \boldsymbol{\varphi}'\right)^{'}, \tag{30}$$

Presuming a diffuse prior (Zellner, 1971:242) as:

$$p\left(\boldsymbol{\beta},\ \boldsymbol{\Sigma^{-1}}\right) = p\left(\boldsymbol{\beta}\right) p\left(\boldsymbol{\Sigma}^{-1}\right) \propto |\boldsymbol{\Sigma}^{-1}|^{\frac{-(F+1)}{2}} \tag{31}$$

the conditional posterior distributions for $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$ can be stated as:

$$p\left(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\Sigma}\right) \sim MVN\left(\left(\boldsymbol{\Sigma^{-1}} \otimes \boldsymbol{X}'\boldsymbol{X}\right)^{-1}\left(\boldsymbol{\Sigma^{-1}} \otimes \boldsymbol{X}'\right)\boldsymbol{y}^{*}, \left(\boldsymbol{\Sigma^{-1}} \otimes \boldsymbol{X}'\boldsymbol{X}\right)\right) \tag{32}$$

$$p\left(\boldsymbol{\Sigma}|\boldsymbol{y},\boldsymbol{X},\boldsymbol{\beta}\right) \sim IW\left(\tilde{\boldsymbol{e}}'\tilde{\boldsymbol{e}},\ \mathbf{Q}\right) \tag{33}$$

$$\tilde{\boldsymbol{e}} = \begin{pmatrix} u_{1,1} & . & . & . & u_{F,1} & v_{1,1} & . & . & . & v_{F,1} \\ . & & & & & & & & & . \\ . & & & & & & & & & . \\ . & & & & & & & & & . \\ u_{1,Q} & . & . & . & u_{F,Q} & v_{1,Q} & . & . & . & v_{F,Q} \end{pmatrix} \tag{34}$$

As earlier stated, the Slutsky matrix should be negative semi-definite to achieve concavity condition (see equation 20). Negativity condition is achieved, in practice, by specifying an accept/reject step in the estimation algorithm such that only the draws ($\boldsymbol{\beta}$) from the distribution in equation 32 which fulfill the negativity condition are kept in the sample employed for inference.

It is important to note that the dependent variable ($\boldsymbol{y^{*}}$) in equation 29 contains both latent and observed elements. All the elements of the vector of continuous (dependent) variables $\mathbf{d}^{*}$ determining the binary outcomes and part of the elements of $\boldsymbol{w^{*}}$ associated with zero expenditures are latent (unobserved). Latency in both instances are regarded as missing data in this study. We employ data augmentation strategy introduced by Tanner and Wong (1987) to handle the missing (incomplete) data problems. Augmenting observed (incomplete) data with latent data ensures unbiased estimation (Tanner and Wong, 1987). (Albert and Chib, 1993) demonstrate how to achieve data augmentation using the Gibbs sampler. They show that latent data can be treated as additional blocks of unknown parameters in the estimation algorithm if the conditional distributions of the latent data can be achieved. In our application, the Gibbs sampler was used to make iterative draws from the conditional distributions of the latent data.

We now turn to the derivation of the conditional distributions. First, we consider the conditional posterior for the missing data. Given the assumption that observations for a particular household are independent of the other households, we introduce the conditional distributions by defining $\boldsymbol{y_q^*}$ and $\hat{\boldsymbol{y_q^*}}$ to include exclusively the elements of $\boldsymbol{y^*}$ and $\hat{\boldsymbol{y^*}} = \boldsymbol{X\beta}$ which correspond to the respective qth household and draw the latent data on household basis. Drawing latent data on food group/subgroup basis is also less difficult. Hence, specifying the precision matrix $\omega = \boldsymbol{\Sigma}^{-1}$, The conditional mean $(\bar{\wedge}_{fq})$ for each of the elements of $\boldsymbol{y^*}$ can be specified as:

$$\bar{\wedge}_{fq} = \hat{y^*_{fq}} + \boldsymbol{\Sigma}_f \boldsymbol{\Sigma}_{-f}^{-1}(\mathbf{y}^*_{-f,q} - \hat{\mathbf{y}}^*_{-f,q}) = \hat{y^*_{fq}} - \omega_{ff}^{-1}\boldsymbol{\omega}_{-f}(\mathbf{y}^*_{-f,q} - \hat{\mathbf{y}}^*_{-f,q}) \tag{35}$$

$$= \hat{y^*_{fq}} - \omega_{ff}^{-1}\sum_{f\neq g}\omega_{gf}(\mathbf{y}^*_{-f,q} - \hat{\mathbf{y}}^*_{-f,q}) \tag{36}$$

and the associated variance (ϑ_f) expressed as:

$$\vartheta_f = \Sigma_{ff} - \boldsymbol{\Sigma_f \Sigma_{-f}^{-1} \Sigma'_{-f}} = \boldsymbol{\omega}_{-f}^{-1} \tag{37}$$

where $\boldsymbol{\omega} = \boldsymbol{\Sigma}^{-1}$ is the precision matrix, Σ_{ff} is the fth on-diagonal element of matrix $(\boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}_f$ is the fth row of $\boldsymbol{\Sigma}$ other than Σ_{ff}, and $\boldsymbol{\Sigma}_{-f}$ denotes the matrix within $\boldsymbol{\Sigma}$ after omitting both the fth column and fth row. ω_{ff} and $\boldsymbol{\omega}_{-f}$ are specified in the same way as Σ_{ff} and $\boldsymbol{\Sigma}_{-f}$. $\hat{\mathbf{y}}_{fq}$ represents the fitted value of y_{fq} corresponding to the qth household. $\hat{d}_{-f,q}$ and $d_{-f,q}$ are vectors within $\hat{\boldsymbol{d}}^*_{\boldsymbol{q}}$ and $\boldsymbol{d}^*_{\boldsymbol{q}}$ respectively, with their fth elements omitted. Following Tiffin and Arnoult (2010), the missing data in the probit equations have conditional distributions of the form:

$$d_{fq} = 0 \quad for\, d^*_{fq}|\boldsymbol{d}^*_{-f,q}, \beta, \mathbf{X}, \Sigma \sim \boldsymbol{N}(\bar{\wedge}_{fq},\, \vartheta_f) I_{[-\infty,0]} \quad \forall_{f,q} \tag{38}$$

$$d_{fq} = 1 \quad for\, d^*_{fq}|\boldsymbol{d}^*_{-f,q}, \beta, \mathbf{X}, \Sigma \sim \boldsymbol{N}(\bar{\wedge}_{fq},\, \vartheta_f) I_{[0,\infty]} \quad \forall_{f,q} \tag{39}$$

and in the demand (share) equations, the form is:

$$w_{fq} = 0 \quad for\, w^*_{fq}|\boldsymbol{d}^*_{-f,q}, \Theta, \mathbf{X}, \Sigma \sim \boldsymbol{N}(\bar{\wedge}_{fq},\, \vartheta_f) I_{[0,1]} \quad \forall_{f\notin C,q} \tag{40}$$

where $I_{[-\infty,0]}$ is an indicator variable equalling one if $d^*_{fq} \in [-\infty, 0]$ and zero otherwise. $I_{[0,\infty]}$ and $I_{[0,1]}$ are similarly specified for d^*_{fq} and w^*_{fq} on their corresponding intervals.

The next issue for discussion pertains to the identification of the probit equations. This is accomplished by putting restriction on the variance-covariance matrix:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix} \tag{41}$$

such that Σ_{vv}=**I**. Where **I** is a F x F identity matrix. Consequently, rather than using equation 33 for drawing $\boldsymbol{\Sigma}$, the conditional posterior for each sub-matrix in $\boldsymbol{\Sigma}$ was obtained. The process begins with the specification of Q x F matrices:

$$\tilde{\boldsymbol{u}} = \begin{pmatrix} u_{11} & . & . & . & u_{F1} \\ . & & & & \\ . & & & & \\ . & & & & \\ u_{1Q} & . & . & . & u_{FQ} \end{pmatrix} \tag{42}$$

$$\tilde{\boldsymbol{v}} = \begin{pmatrix} v_{11} & . & . & . & v_{F1} \\ & & & & . \\ & & & & . \\ & & & & . \\ v_{1Q} & . & . & . & v_{FQ} \end{pmatrix} \tag{43}$$

It follows from the multivariate normal properties (Geweke, 2005:171) that the conditional mean and variance can be stated as:

$$E(\mathbf{u}|\mathbf{v}) = \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vv}^{-1}\mathbf{v} \tag{44}$$

$$E(\mathbf{u}'\mathbf{u}|\mathbf{v}) = \boldsymbol{\Sigma}_c = \boldsymbol{\Sigma}_{uu} - \boldsymbol{\Sigma}_{uv}\boldsymbol{\Sigma}_{vv}^{-1}\boldsymbol{\Sigma}_{vu} \tag{45}$$

where equation 44 is a regression of u on v. Following from the restriction that Σ_{vv}=**I**, the variance-covariance matrix ($\boldsymbol{\Sigma}$) can be parameterised as:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{uu} & \Sigma_{uv} \\ \Sigma_{vu} & \Sigma_{vv} \end{bmatrix} = \begin{pmatrix} \Sigma_c + \rho\rho' & \rho \\ \rho' & \mathbf{I} \end{pmatrix} \tag{46}$$

where ρ=Σ_{uv}. Acknowledging from equations 44 and 45, coupled with the assumption that Σ_{vv}=**I**; ρ and Σ_c are thus, the vector of coefficients and covariance matrix of disturbance term ($\boldsymbol{\varsigma}$), respectively, in the SUR stated below:

$$\mathbf{u} = \boldsymbol{\rho}\mathbf{v} + \boldsymbol{\varsigma} \tag{47}$$

where $\boldsymbol{\varsigma}\sim$ N(0, $\boldsymbol{\Sigma}_c$). Assuming a diffuse prior, the conditional posteriors for $\boldsymbol{\rho}$ and Σ_c can be specified as:

$$\boldsymbol{\rho}|\boldsymbol{\Sigma}_c^{-1} \sim N\left[\left(\tilde{\mathbf{u}}\boldsymbol{\Sigma}_c^{-1}\tilde{\mathbf{u}}\right)^{-1}\bar{\mathbf{u}}'\bar{\mathbf{v}}, \left(\tilde{\mathbf{u}}\boldsymbol{\Sigma}_c^{-1}\tilde{\mathbf{u}}\right)^{-1}\right] \tag{48}$$

and

$$\boldsymbol{\Sigma}_c^{-1}|\varrho \sim IW(\boldsymbol{\varsigma}'\boldsymbol{\varsigma}, \mathbf{Q}) \tag{49}$$

respectively. $\boldsymbol{\rho}$ and Σ_c are employed in the relations indicated in equation 46 as the groundwork for sampling the restricted variance-covariance matrix (equation 46). The estimation algorithm is stipulated in subsection 3.1.

3.1 Estimation (Gibbs) Algorithm

The estimation algorithm describes the key steps involved in estimating the parameters of the model. The estimation algorithm are systematically presented below.

1. Assume (select) the starting values for $\boldsymbol{y}^*$ and $\boldsymbol{\Sigma}$.

2. Employ the newly drawn values of $\boldsymbol{y}^*$ obtained from Steps 4 and 5 and $\boldsymbol{\Sigma}$ generated from Step 6 or those presumed in Step 1 (in case this is the first draw) to draw the vector of parameters ($\boldsymbol{\beta}$).

3. Calculate the Slutsky matrix from equation 20 using the appropriate elements of $\boldsymbol{\beta}$ and check if the matrix is negative semi-definite. If so, include the drawn (sampled) vector of parameters ($\boldsymbol{\beta}$). Otherwise, go back to the former (immediate past) draw of $\boldsymbol{\beta}$.

4. With the vectors of parameters ($\boldsymbol{\beta}$) generated in Step 2, compute $\hat{\boldsymbol{y}}^* = \boldsymbol{X}\boldsymbol{\beta}$. Employ the appropriate elements in $\hat{\boldsymbol{y}}^*$ and $\boldsymbol{\Sigma}$ from step 6 or the assumed starting values in Step 1 (in case this is the first draw) to estimate the mean as well as the variance of conditional distributions from equations 36 and 37.

5 (a) To obtain the latent (missing) data for the probit equations, apply the appropriate mean and variance obtained from Step 4 to draw from the truncated normal distributions specified in equations 38 and 39.

5 (b) To generate latent data for the demand (share) equations where zero expenditures are reported, utilise the appropriate mean and variance obtained from Step 4 to take draws from the truncated normal distributions in equations 40.

6. Use $\boldsymbol{\beta}$ obtained from Step 2 and $\boldsymbol{y}^*$ from Steps 4 and 5 to draw covariance matrix $\boldsymbol{\Sigma}$.

7. Draw ρ from the normal distribution in equation 48

8. Draw $\boldsymbol{\Sigma}_c$ from the inverted Wishart distribution in equation 49.

9. Use ρ and $\boldsymbol{\Sigma}_c$ to complete the covariance matrix $\boldsymbol{\Sigma}$ in equation 46.

10. Return to Step 2.

4 Description of Data

The Nigeria Living Standard Survey (NLSS) 2003/2004 household data and the food price data for 2003/2004 obtained from the National Bureau of Statistics, Nigeria are employed for this study. The NLSS data is the largest, most comprehensive available micro-data in the country which can be used for demand analysis. The NLSS data were collected at the household levels from rural and urban sectors in various enumeration areas across the 36 states of the country including the federal capital territory (Abuja) through a two-stage stratified sampling technique from September, 2003 to August 2004. A total of 19158 households were sampled by means of questionnaire. The price data were collected on a monthly for the year 2003 and 2004 in both rural and urban sectors across the 36 states of the country including Abuja and in most of the enumeration areas covered by the household survey (NLSS). The food items in the food price file correspond with those in the NLSS data. There are 70 food items in the price file excluding tobacco.

The NLSS household data were compiled in different files having names that suitably describe the items contained in them. Although data collected covers different aspects of household livelihoods, data on food purchased in the markets (from food purchase file), quantities of food consumed out of what the household produce (from own-consumption file) and data from the the household expenditures and socio-demographic characteristics files are relevant to the study. Food data were collected from each household on a weekly basis over a period of six consecutive weeks. The values of foods purchased in the markets were recorded in the Naria (national currency) while the quantities of foods consumed from household production were reported either in the standard or local measures. There are effectively 133 food items listed in the food files besides tobacco and cigarettes. However, not all household reported consumption of the entire food items. The own-consumption file also contains information on the amount (price) each of the "own-consumed" food items could be sold for in the market. These prices are referred to as proxy prices in this study. The food quantities were converted to kilogramme and the associated prices expressed in price/kilogramme.

Relevant data in the household expenditure and demographic characteristics file include expenditure on non-food items and non-food price index; age, gender and educational status of household head; the sector (rural/urban) and the geopolitical zone where household belongs; Food and Agriculture

Organization (FAO) adult equivalent (AE) household size and dates of first visit to each household. Each of the NLSS household data files contains information on the states, sectors, enumeration areas and household numbers (unique identifiers). The NLSS household data files were first merged together based on these information. There are 18880 households that merged from all the household data files. The resultant bigger file is referred to as household data file.

Thereafter, the food price data were merged with the already merged household data filed based on state, sectors, enumeration areas and month and year of data collection. Households in a given enumeration area are presumed to be facing similar market price situation. Only 13950 households of entire households eventually merged successfully with the food price data as some enumeration areas and months in the price file do not merge perfectly with the those in the NLSS household data file. Where a food item is consumed and the price is not available in the food price file, the average of the proxy price reported of the particular food items reported by household(s) in the corresponding enumeration area, sector or state is computed and applied as the market price. Estimated expenditures on own-consumed foods were computed and added to the corresponding food expenditure data from the food purchase file. Expenditures on food and non food items were converted to monthly information and discounted by the adult equivalent household size. The per capita adult equivalent information are used for analysis. Households with per capita calorie intake below 500 kcal per day were removed from the 13950 households as hardly would an average adult survive with that daily intake level. A total of 13142 households were finally employed for analysis.

For the purpose of analysis, we first partition items into two categories-food and non-food items. The entire food entire were also divided into six groups (staples, animal products, beverages, fruits vegetables, fats and oils and seeds and nuts). Since our focus is on staple, we limit further partitioning to staples. We classified staples into four food subgroups namely: cereals, beans, tubers and snacks. Classification of foods items are based on previous studies on food demand and nutrition in the country (Oguntona and Akinyele, 1995; Maziya-Dixon et al., 2004; Obayelu, Okoruwa and Ajani, 2009). The proportion of households having zero consumption records on cereals, beans, tubers and snacks is approximately 1.33, 24, 20.33 and 77.66 per cent respectively. We constructed the Thornqvist Theil price index to represent the price of each food aggregates.

We estimated demand models at hierarchy of three budgeting stages. At the first stage, we estimated a single demand (share) equation with aggregate food budget share as dependent variable. The second stage featured estimation of demand subsystem for the six food groups while the third stage involved estimation of a demand subsystem for the four staple food subgroups. We employ the appropriate elasticities obtained from the estimated demand models to compute unconditional price and total expenditure elasticities for the four staple food subgroups following Edgerton (1997). Households were grouped into five income quintiles in order to explore differences in household demand responses.The demographic variables included in the demand equations are presented in Table 1.

Table 1: Demographic variables (factors) included in the demand (share) equations

Demographic variables	Description of Variables
Household composition	Adults Only^a; Family with children and two adults; Family with adolescents and two adults Family with children and more than two adults; Family with adolescents and more than two adults Family without children and adolescents and more than two adults.
Age composition of household Head	Less than 31 years^a; 30 to 45 years; 46-60 years, Above 60 years
Educational status of household Head (Highest education level attained)	No formal/primary school education^a; secondary school education, College of Education (NCE)/National Diploma (ND); Higher National Diploma (HND)/University Education.
Geo-political zones	South West zone^a; South South zone; South East zone; North Central zone; North East zone; North West zone
Sector	Urban^a; Rural
Gender of the household head	Male^a; Female
Received to credit (former/ informal sources)	Not receive^a; Received
Received to remittance	Not received^a; Received
Seasons of survey	Hungry/lean season^a; Harvest/Surplus season

All variables are computed as dummies. ^{a} Variable that is dropped in each demographic category. Omitted variables serve as the reference groups for result's interpretation.

5 Results and Discussion

The results of the unconditional price and total expenditure elasticities for cereals, beans, tubers and snacks are presented in Table 2. We do not report the results of the compensated and uncompensated elasticities obtained at each stage of the hierarchy in this paper. They are available on request. However, we present the results of demographic factors influencing demand for staple food subgroups (Table 3). All the parameters reported are the mean values generated from the Gibbs sample. For brevity, we present results for the lowest, middle (third) and the highest income quintiles.

Table 2 shows that all own price elasticities have the expected negative signs. This is a result of the negativity constraint imposed on the curvature of the expenditure function. Snacks has the largest own-price elasticities in each of the income quintiles; suggesting that demand for snacks is the most sensitive to changes in own-price. Own-price elasticities of cereals are consistently high in the lowest, middle and highest income quintiles with values -0.938, -0.926 and -1.018 respectively. This indicates that households in the highest income quintile are the most sensitive to changes in cereal prices. Although demand for cereals is own-price inelastic in the lowest and the middle income quintiles, elasticity values are close to unity. Demand for cereals is own-price elastic in the highest income quintile.

Cross-price elasticities indicate a mix of gross substitutability and complementarity relationships among staple food subgroups; indicating the role of total expenditure (income) effects in stimulating cross-price responsiveness to food consumption among households. The complementary relationships found between tubers and beans in the lowest and middle income quintiles are expected. Households in Nigeria consume tubers such as yam and cocoyam with beans in form of porridge. Consumption of tubers and tuber products with bean cakes/meals and bean soups are also common, especially in the southern parts of the country. Total expenditure (income) elasticities on food subgroups decline with higher income levels in line with the Engel's, law. All staple food subgroups are normal goods across income quintiles except for snack food subgroup which appears as inferior good in the highest income quintile. Cereals and tubers have the highest total expenditure elasticities and are luxury goods among households in the lowest income quintile. This suggests that an average household in the poorest segment of the population will increase consumption of tubers and cereals more than other staple foods if household income improves.

We now consider the influence of demographic factors on demand for staple food subgroups (Table 3). We indicate a significant variable with an asterisk; showing that the 95 percentiles of its highest HPDIs excludes zero. The results of snacks subgroup are excluded as none of the demographic variables included in the model had significant effects on its demand. The results indicate that rural households in the middle and the highest income quintiles consume more tubers and less beans than there urban counterparts. Variations in demand for staple food subgroups at the zonal levels is also well pronounced. For instance, while consumption of cereals is higher in the North Central, North East and North West zones, consumption is lower in the South South and South East zones. The opposite is the case for tubers. Presence of children and adolescents in the households leads to high demand for tubers. Demand for tubers is also higher among household headed by older people compared to households headed by people below 31 years.

6 Conclusions

A major challenge confronting applied econometricians is how to suitably model the zero-valued observations (censoring) in the dependent variables while estimating demand systems using microdata. In this paper, we applied a Bayesian approach to estimate household food demand systems using a multivariate double-hurdle model to handle zero observations attributable to non-participation decisions. Demand patterns feature a mix of gross substitutability and complementarity among staple food subgroups, indicating the importance of total expenditure(income) effect in determining cross-price responsiveness. Presence of children and adolescents, rural-urban and zonal (regional) variations and age of household head, among others, have significant influence on demand for cereals, beans and tubers. Total expenditure (income) elasticities on food subgroups decline with higher income levels.

Although cereals, beans and tubers are normal goods, they are all necessary goods in the middle and highest income quintiles. Total expenditure elasticities of cereals and tubers greater than unity in the lowest income quintile. The implication is that improved income would lead to higher consumption of cereals and tubers among households in the poorest income quintile. In sum, while income growth would enhance consumption of staples among households generally, targeted interventions such as cash or food stamp transfers are advocated to boost access to major staple foods among the poorest household groups in the country.

Table 2: Unconditional elasticities for staple food subgroups

Quantity	Price				Total expenditure
	Cereals	Beans	Tubers	Snacks	elasticity
			Lowest income quintile		
Cereals	-0.938	0.018	0.041	-0.054	1.039
Beans	0.419	-0.968	-0.093	0.115	0.586
Tubers	0.060	-0.080	-0.915	-0.029	1.073
Snacks	1.623	1.765	1.477	-5.475	0.677
			Middle income quintile		
Cereals	-0.926	0.076	0.041	0.072	0.806
Beans	0.502	-0.847	0.034	-0.145	0.498
Tubers	-0.008	-0.024	-0.799	-0.010	0.919
Snacks	1.101	-0.290	0.197	-1.509	0.547
			Highest income quintile		
Cereals	-1.018	0.083	0.257	0.079	0.184
Beans	0.365	-0.928	0.125	-0.047	0.149
Tubers	0.219	0.015	-0.835	-0.065	0.204
Snacks	7.544	-0.414	-2.667	-3.667	-0.245

Table 3: Demographic factors affecting demand for staple food subgroups

			Age of Household head		Classification by geo-political zones					Household composition				
	Sector	A-Credit	46-60years	> 60years	S-South	S-East	N-Central	N-East	N-West	FWC2A	FWA2A	FWC>2A	FWA>2A	NCA>2A
							Lowest income quintile							
Cereals	0.001	-0.028	0.012	-0.009	-0.063*	-0.024	0.134*	0.278*	0.251*	-0.006	0.001	-0.004	-0.003	-0.012
Beans	-0.012	-0.005	0.007	0.003	-0.028*	-0.031*	-0.036*	-0.057*	-0.013	0.003	0.005	0.000	0.002	-0.005
Tubers	0.005	0.025	-0.017	0.016	0.081*	0.047	-0.087*	-0.219*	-0.240*	0.003	0.006	0.013	0.004	0.020
							Middle income quintile							
Cereals	-0.014	-0.025	0.001	-0.010	-0.080*	-0.045*	0.099*	0.260*	0.271*	-0.001	-0.004	0.003	-0.021	-0.020
Beans	-0.022*	-0.002	-0.006	-0.003	-0.022*	-0.028*	-0.034*	-0.039*	-0.009	-0.010	-0.003	-0.005	-0.007	-0.009
Tubers	0.021*	0.033*	0.035*	0.034*	0.087*	0.078*	-0.062*	-0.208*	-0.236*	0.029*	0.024*	0.027*	0.036*	0.059*
							Highest income quintile							
Cereals	-0.024*	-0.012	-0.040*	-0.031*	-0.059	-0.036*	0.127*	0.278*	0.261*	-0.005	-0.023*	-0.015	-0.002	0.006
Beans	-0.016*	-0.007	-0.005	-0.004	-0.011	-0.019*	-0.030*	-0.040*	-0.028*	-0.010*	-0.006	-0.007	-0.007	-0.018*
Tubers	0.026*	0.023	0.033*	0.019	0.066*	0.050*	-0.067*	-0.221*	-0.219*	0.024*	0.008	0.027	0.011	-0.022

Note: A-Credit=Access to credit, >=Greater than or equal to, S-South=South South zone, S-East=South East zone, N-Central=North Central zone, N-East=North East, N-West = North West zone, FWC2A = Family with children and two adults, FWA2A=Family with adolescents and two adults, FWC>2A=Family with children and more than two adults, FWA>2A=Family with adolescents and more than two adults, NCA>2A=Family without children and adolescents and more than two adults, * Significant variable with 95 percentile HPDIs excluding zero. The snacks staple food subgroup is excluded as none of the associated demographic variables is significant in the income quintiles.

References

Adejobi, A. O. and Babatunde, R. O. (2010). Analysing the level of market orientation among rural farming households in northern Nigeria, *African Journal of General Agriculture* **6(4)**: 255–261.

Akinbode, S. O. and Dipeolu (2012). Double-hurdle model of fresh fish consumption among urban households in South-West Nigeria, *Current Research Journal of Social Sciences* **4(6)**: 431–439.

Akinyele, I. O. (2009). Ensuring food and nutrition security in rural Nigeria: an assessment of the challenges, information needs, and analytical capacity. Nigeria Strategy Support Program (NSSP) Background paper 7, International Food Policy and Research Institute.

Albert, J. A. and Chib, S. (1993). Bayesian analysis of binary and polychotomous response data, *Journal of the American Statistical Association* **88(422)**: 669–679.

Aliyu, A. A., Oguntunde, O. O., Dahiru, T. and Raji, T. (2012). Prevalence and determinants of malnutrition among pre-school children in northern Nigeria, *Pakistan Journal of Nutrition* **11(11)**: 1092–1095.

Amemiya, T. (1974). Multivariate regression and simultaneous equation models when the dependent variables are truncated normal, *Econometrica* **42**: 999–1012.

Angulo, A. M., Gil, J. M. and Gracia, A. (2001). The demand fror alcoholic beverages in Spain, *Agricultural Economics* **26(2001)**: 71–83.

Arene, C. J. and Anyaeji, R. C. (2010). Determinants of food security among households in nsukka metropolis of enugu state, nigeria, *Pakistan Journal of Social Sciences* **30(1)**: 9–16.

Ashagidigbi, W. M., Yusuf, S. A. and Okoruwa, V. O. (2012). Determinants of households' food demand in nigeria, *World Rural Observations* **4(4)**: 17–28.

Blisard, N. and Blaylock, J. (1993). Distinguishing between market participation and infrequency of purchase models of butter demand, *American Journal of Agricultural Economics* **75**: 314–320.

Cragg, J. G. (1971). Some statistical models for limited dependent variables with application to the demand for durables, *Econometrica* **39(5)**: 829–844.

Deaton, A. and Irish, M. (1984). Statistical models for zero expenditures in household budgets, *Journal of Public Economics* **23(1-2)**: 59–80.

Dong, D. D., Gould, B. W. and Kaiser, H. M. (2004). Food demand in mexico: An application of the ameniya-tobin approach to the estimation of a censored food system, *American Journal of Agricultural Economics* **86(4)**: 1094–1107.

Edgerton, D. L. (1997). Weak separability and the estimation of elasticity in multistage demand system, *American Journal of Agricultural Economics* **79(1)**: 62–79.

FAO (2012). The state of food insecurity in the world 2012. Rome , (Available at http://www.fao.org/docrep/016/i3027e/i3027e.pdf).

Fregene, T. B. and Bolorunduro, P. I. (2009). Role of women in food security and seasonal variation of expenditure pattern in coastal fishing communities in Lagos State, *Journal of Agricultural Extension* **13 (2)**: 21–33.

Gegios, A., Amthor, R., Maziya-Dixon, B., Egesi, C., Mallowa, S., Nungo, R., Gichuki, S., Mbanaso, A. and Manary, M. J. (2010). Children consuming cassava as a staple food are at risk for inadequate zinc, iron, and vitamin a intake, *Plant Foods for Human Nutrition* **65(1)**: 64–70.

Geweke, J. (2005). *Contemporary Bayesian Econometrics and Statistics*, New Jersey, USA, Wiley & Son Inc.

Goon, D. T., Toriola, A. L., Shaw, B. S., Amusa, L. O., Monyeki, M. A., Akinyemi, O. and Alabi, O. A. (2011). Anthropometrically determined nutritional status of urban primary schoolchildren in Makurdi, Nigeria, *BMC Public Health* **11:769**: 1–8.

Kasteridis, P. and Yen, S. (2012). U.S. demand for organic and conventional vegetables: a Bayesian censored system approach, *Australian Journal of Agricultural and Resource Economics* **56(3)**: 405–425.

Kimhi, A. (1999). Double-hurdle and purchase-infrequency demand analysis: a feasible integrated approach, *European Review of Agricultural Economics* **26(4)**: 425–442.

Maki, A. and Garner, T. I. (2004). The gap between macro and micro economic statistics: Estimation of the misreporting model using micro-data sets, Available at http://repec.org/esAUSM04/up.17601.1075102278.pdf.

Maki, A. and Nishiyama, B. (1996). An analysis of under-reporting for micro-data sets: The misreporting or double-hurdle model, *Economics Letters* **52(3)**: 211–220.

Maziya-Dixon, B. M., Akinyele, I. O., Oguntona, E. B., Nokoe, S., Sanusi, R. A. and Harris, E. (2004). *Nigeria food consumption and nutrition survey 2001-2003 Summary*, International Institute of Tropical Agriculture (IITA).

Musa, U., Hati, S. S. and Mustapha, A. (2012). Levels of fe and zn in staple cereals: Micronutrient deficiency implications in rural northeast Nigeria, *Food and Public Health* **2(2)**: 28–33.

NBS (2012). Consumption pattern in nigeria 2009/10: Preliminary report (2012), (Available at http://www.nigerianstat.gov.ng/).

Newman, C., Henchion, C. and Matthews, A. (2001). Infrequency of purchase and double-hurdle models of irish households' meat expenditure, *European Review of Agricultural Economics* **28(4)**: 393–412.

Obayelu, A. E. (2010). Global food price increases and nutritional status of Nigerians: the determinants, coping strategies, policy responses and implications, *ARPN Journal of Agricultural and Biological Science* **5(2)**: 67–80.

Obayelu, A. E., Okoruwa, V. O. and Ajani, O. I. Y. (2009). Cross-sectional analysis of food demand in the north central, nigeria: The quadratic almost ideal demand system (quaids) approach, *China Agricultural Economic Review* **1(2)**: 173–193.

Obayelu, A. E., Okoruwa, V. O. and Oni, O. A. (2009). Analysis of rural and urban households' food consumption differential in the north-central, nigeria: A micro-econometric approach, *Journal of Development and Agricultural Economics* **1(2)**: 18–26.

Ogunniyi, L. T., Ajao, A. O. and Oladejo, J. A. (2012). Food consumption patterns in ogbomoso metropolis of Oyo State, Nigeria, *Journal of Agriculture and Social Research* **12(1)**: 74–83.

Oguntona, E. B. and Akinyele, I. O. (1995). *Nutrient Composition of Commonly Eaten Food in Nigeria -Raw, Processed and Prepared*, Food Basket Foundation, Ibadan.

Olagunju, F. I., Oke, J. T. O., Babatunde, R. and Ajiboye, A. (2012). Determinants of food insecurity in ogbomoso metropolis of oyo state, nigeria, *PAT* **8 (1)**: 111–124.

Olayemi, A. O. (2012). Effects of family size on household food security in Osun State, Nigeria, *Asian Journal of Agriculture and Rural Development* **2(2)**: 136–141.

Omuemu, V. O., Otasowie, E. M. and Onyiriuka, U. (2012). Prevalence of food insecurity in Egor local government area of Edo State, Nigeria, *Annals of African Medicine* **11(3)**: 139–145.

Orewa, S. I. and Iyangbe, C. (2010). The struggle against hunger: the victims and the food security strategies adopted in adverse conditions, *World Journal of Agricultural Sciences* **6(6)**: 740–745.

Perali, F. and Chavas, J. (2000). Estimation of censored demand equations from large cross-section data, *American Journal of Agricultural Economics* **82(4)**: 1022–1037.

Tanner, M. A. and Wong, W. H. (1987). The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association* **82(398)**: 528–540.

Tiffin, R. and Arnoult, M. (2010). The demand for a healthy diet: estimating the almost ideal demand system with infrequency of purchase, *European Review of Agricultural Economics* **37(4)**: 501–521.

Tobin, J. (1958). Estimation of relationships for limited dependent variables, *Econometrica* **26**: 24–36.

Ubesie, A. C., Ibeziako, N. S., Ndiokwelu, C. I., Uzoka, C. M. and Nwafor, C. A. (2012). Under-five protein energy malnutrition admitted at the University of Nigeria Teaching Hospital, Enugu: a 10 year retrospective review, *Nutrition Journal* **11(43)**: 1–7.

Yen, S. T. and Jensen, H. H. (1996). Determinants of household expenditures on alcohol, *Journal of Consumer Affairs, Summer* **30(1)**: 48–67.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.*