Heterogeneity of Household Food Expenditure Patterns in South Africa

Lesiba Bopape*

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Long Beach, California, July 23-26, 2006

---

* Ph.D. candidate, Department of Agricultural Economics, Michigan State University, East Lansing, MI 48823. I would like to thank Robert Myers and Jeffrey Wooldridge for their helpful comments and guidance. All remaining errors are my own.

The Heterogeneity of Household Food Expenditure Patterns in South Africa

1. Introduction

Aggregate per capita availability data suggest that South Africa is food secure in almost all basic foodstuffs. Furthermore, South Africa has the highest per capita income in Sub-Saharan Africa, and is categorized as a middle-income country with average per capita gross national income of US $3,650 in 2004 (World Bank, 2004). These facts suggest that hunger and food security should not be major policy issues in the country. However, these aggregate data mask a highly unequal distribution of income and a huge divide between relatively affluent urban areas and destitute conditions in many rural communities. The richest 20% of the population receives over 60% of the income while the poorest 20% receives less than 3% (World Development Report, 2002). At the household level, over 30% of the population is categorized as vulnerable to food insecurity and over 20% of the children are estimated to be stunted and vitamin A deficient (Human Science Research Council, 2004).

Policies designed to reduce income inequality, hunger, and malnutrition have had mixed results. Major social, economic, and political reforms introduced since the demise of apartheid and the emergence of democratic government in 1994 clearly have redistributed wealth. But income inequality and household food insecurity remain. One of the problems is that little is known about how food expenditure patterns differ across different income groups, and across different geographic regions. Without a thorough understanding of the heterogeneity of food expenditure patterns, and how these patterns are changing over time, it will continue to be difficult to design policies that improve food security effectively over a broad range of heterogenous low-income households.

This paper seeks to improve knowledge and understanding of the heterogeneity in food expenditure patterns in South Africa. The study makes use of an unusually rich panel dataset on household food consumption, collected as part of the KwaZulu-Natal Income Dynamics Study (KIDS). The KIDS dataset contains detailed information on household socioeconomic and demographic characteristics, which permit heterogeneity effects to be analyzed. The dataset followed the same households over a ten-year period, with surveys in 1993, 1998, and 2004, to study changes in their incomes and expenditure patterns. Data on prices of various food products consumed by households were also collected.

This paper uses the KIDS data to estimate demand functions for seven food groups— grains, meat and fish, fruits and vegetables, dairy, oils and fats, sugar, and all other foods. The paper also examines how food expenditure patterns differ between rural and urban households, as well as across income groups. The Quadratic Almost Ideal Demand System (QUAIDS) of Banks *et al*. (1997) is used to estimate price and expenditure elasticities, as well as the impact of household demographic characteristics on food expenditure patterns. There are two main motivations for this study. First, there is no previous known application of the QUAIDS model to food expenditure data in South Africa. Also, most previous food demand studies ignore the fact that expenditure may be endogenous in budget share equations. This study explicitly tests and controls for expenditure endogeneity. Expenditure endogeneity is controlled for using an augmented regression approach suggested by Blundell and Robin (1999). Secondly, this study builds on the work of Banks *et al*. (1997) to develop a Lagrange Multiplier (LM) test that can be used to determine whether the demand model should be specified with a quadratic (QUAIDS) or a linear (AIDS) expenditure term. The usefulness of this test is that it can be conducted without having to explicitly estimate the highly nonlinear QUAIDS model. No other study was found to have explicitly conducted this test, certainly not with South African data.

The rest of the paper is organized as follows. Section 2 presents the demand model and provides a discussion of econometric approaches to testing for the quadratic expenditure specification and expenditure endogeneity. Section 3 describes the data, while section 4 presents the empirical results. Section 5 summarizes the main findings and concludes.

2. Theoretical Framework

The almost ideal demand system (AIDS) of Deaton and Muellbauer (1980) has been a popular functional form to model demand behavior. The popularity of the AIDS model is due to its many desirable properties, particularly the facts that it satisfies the axioms of choice exactly and allows for consistent aggregation of individual demands to market demands. The AIDS model is a member of the Price-Independent Generalized Logarithmic (PIGLOG) class of demand models (Muellbauer, 1976), which have budget shares that are linear functions of log total expenditure. However, there is increasing evidence that further terms in total expenditure may be required for at least some of the budget share equations (Lewbel, 1991; Blundell *et al*., 1992). Furthermore, by allowing only the linear expenditure term, the AIDS model makes the restrictive assumption that expenditure elasticities are constant at all expenditure levels. A natural alternative to the AIDS model would be a more general model that includes further terms in the expenditure variable, thus allowing expenditure elasticities to differ with the level of expenditure level (e.g., allowing goods to be luxuries at some expenditure levels and necessities at others). Banks *et al*. (1997) develop an extension of the AIDS model, the quadratic almost ideal demand system (QUAIDS), which is quadratic in log total expenditure.

The QUAIDS model is a generalization of PIGLOG preferences based on the following indirect utility (V) function:

$$\ln V = \left\{ \left[\frac{\ln x - \ln a(\mathbf{p})}{b(\mathbf{p})} \right]^{-1} + \lambda(\mathbf{p}) \right\}^{-1} \tag{1}$$

where x is total expenditure, **p** is a vector of prices, $a(\mathbf{p})$ is a function that is homogenous of degree one in prices, and $b(\mathbf{p})$ and $\lambda(\mathbf{p})$ are functions that are homogeneous of degree zero in prices. As in the original AIDS model, ln $a(\mathbf{p})$ and ln $b(\mathbf{p})$ are specified as the translog and Cobb-Douglas equations:

$$\ln a(\mathbf{p}) = \alpha_0 + \sum_{i=1}^{K} \alpha_i \ln p_i + \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{K} \gamma_{ij} \ln p_i \ln p_j \tag{2}$$

$$b(\mathbf{p}) = \prod_{i=1}^{K} p_i^{\beta_i} \tag{3}$$

where $i = 1, \ldots, K$ denote commodities. The function $\lambda(\mathbf{p})$ is specified as:

$$\lambda(P) = \sum_{i=1}^{K} \lambda_i \ln p_i \qquad \text{where} \qquad \sum_{i=1}^{K} \lambda_i = 0 . \tag{4}$$

Application of Roy's identity to (1) gives the QUAIDS budget share equations. To control for varying preference structures and heterogeneity across households, we incorporate demographic variables ($\mathbf{z}$) into the QUAIDS model through the linear demographic translating method (Pollak and Wales, 1978). This leads to the following empirical specification of the QUAIDS budget share equations:

$$w_i \quad = \quad \alpha_i + \sum_{j=1}^{K} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{x}{a(\mathbf{p})}\right] + \frac{\lambda_i}{b(\mathbf{p})}\left\{\ln\left[\frac{x}{a(\mathbf{p})}\right]\right\}^2 + \sum_{s=1}^{L} \delta_{is} z_s \tag{5}$$

where $\mathbf{z}_{\mathrm{s}} = (z_1, ..., z_L)$ is a set of demographic variables. Formulas for the QUAIDS expenditure and price elasticities are derived by differentiating the budget share equations with respect to ln x and ln p_j, respectively. Following Banks *et al.* (1997), we simplify the expressions for the elasticity formulas by using the intermediate results:

$$\mu_i \equiv \frac{\partial w_i}{\partial \ln x} = \quad \beta_i + \frac{2\lambda_i}{b(\mathbf{p})}\left\{\ln\left[\frac{x}{a(\mathbf{p})}\right]\right\} \tag{6}$$

$$\mu_{ij} \equiv \frac{\partial w_i}{\partial \ln p_j} = \gamma_{ij} - \mu_i\left(\alpha_j + \sum_{l=1}^{K} \gamma_{jl} \ln p_l\right) - \frac{\lambda_i \beta_j}{b(\mathbf{p})}\left\{\ln\left[\frac{x}{a(\mathbf{p})}\right]\right\}^2. \tag{7}$$

In terms of the μ_i, the formula for expenditure elasticities can be written as:

$$e_i = 1 + \frac{\mu_i}{w_i}. \tag{8}$$

The expression for the Marshallian or uncompensated price elasticities can be written as:

$$e_{ij}^u = \frac{\mu_{ij}}{w_i} - \delta_{ij} \tag{9}$$

where δ_{ij} is the Kronecker delta taking the value $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$. The Hicksian or compensated price elasticities are obtained from the Slutsky equation:

$$e_{ij}^c = e_{ij}^u + w_j e_i. \tag{10}$$

The theoretical restrictions of adding-up, homogeneity, and symmetry are imposed on the parameters to ensure integrability of the demand system (Moro and Sckokai, 2000). Addding-up simply requires that the household does not spend more than its total budget (i.e., $\sum_i w_i = 1$), and can be expressed in terms of model parameters as:

$$\sum_{i=1}^{K} \alpha_i = 1 \qquad \sum_{i=1}^{K} \beta_i = 0 \quad \sum_{i=1}^{K} \lambda_i = 0 \quad \sum_{i=1}^{K} \gamma_{ij} = 0 \quad \forall j . \tag{11}$$

Marshallian demands are homogenous of degree zero in $(\mathbf{p}, m)$, and this property is satisfied by imposing the parametric restrictions:

$$\sum_{j=1}^{K} \gamma_{ij} = 0 \quad \forall j . \tag{12}$$

Slusky symmetry requires that

$$\sum_{j=1}^{K} \gamma_{ij} = 0 \quad \forall j . \tag{13}$$

These restrictions (adding-up, homogeneity, and symmetry) are imposed during estimation.

2.1 Quadratic expenditure specification test

Our test for whether the quadratic expenditure term should be included in the demand model— and therefore, whether QUAIDS or AIDS is appropriate for modeling demand behavior— builds on the work by Banks *et al*. (1997). In particular, the implication of corollary 2 in Banks *et al*. (p.533) is that a utility-derived demand system that is rank 3 and exactly aggregable cannot have coefficients on both the linear and the quadratic expenditure terms that are independent of prices.[1] In other words, if a rank 3 exactly aggregable demand system that is derived from utility theory has a coefficient on the linear expenditure term that is independent of prices, then it *must* have a coefficient on the quadratic expenditure term that is price dependent. Since the QUAIDS model, which is rank 3 and exactly aggregable, has a coefficient on the linear expenditure term that is independent of prices, then testing for its (i.e., QUAIDS) functional form involves testing for the statistical significance of prices in the coefficient on the quadratic expenditure term.

To derive the test, it is necessary to relax the theoretical constraint in the QUAIDS model that the β_i parameters in the Cobb-Douglas price aggregator $b(\mathbf{p})$ are the same as the β_i coefficients on the linear expenditure terms (that is, the coefficients on $x/a(\mathbf{p})$).[2] For this purpose, define a new price aggregator $b(\mathbf{p}) = \prod_{i=1}^{K} p_i^{\theta_i}$, where θ_i has been used in place of β_i, and θ_i and β_i are allowed to differ from each other. For ease of exposition, we absorb all the terms not involving the quadratic expenditure term into the vector **q** and their associated parameters (i.e., parameters not involving the θ_i's) into the vector $\boldsymbol{\varphi}$. The expenditure share equations (5) can now be expressed as:

$$w_i \;=\; g_i(\mathbf{q},\boldsymbol{\varphi}) + \lambda_i \left(\prod_{i=1}^{K} p_i^{\theta_i} \right)^{-1} \left\{ \ln\left[\frac{x}{a(\mathbf{p})} \right] \right\}^2 + \varepsilon_i \qquad (14)$$

---

[1] The rank of a demand system is defined as the dimension of the space spanned by its Engel curves (Lewbel, 1991). The rank of demand systems of the form $w_i = A_i(\mathbf{p}) + B_i(\mathbf{p}) \ln x_R + C_i(\mathbf{p}) g(x_R)$, where $x_R = x/a(\mathbf{p})$, and $A_i(\mathbf{p})$, $B_i(\mathbf{p})$, $C_i(\mathbf{p})$ and $g(x_R)$ are differentiable functions, is equal to the $N \times 3$ matrix of Engel curve coefficients, having rows $[A_i(\mathbf{p}):B_i(\mathbf{p}):C_i(\mathbf{p})]$ for goods i (Banks *et al*., 1997). This matrix has three columns, so 3 is the maximum possible rank of the demand system. Exactly aggregable demand systems are defined as demand systems that are linear in functions of x (Banks *et al*., 1997).

[2] This is because if we maintain the restriction that the β_i's in $b(\mathbf{p})$ are the same as the β_i coefficients on the linear expenditure term, then the null hypothesis that the β_i's in $b(\mathbf{p})$ are all zero will make the second term, $\ln(x/a(\mathbf{p}))$, in equation (5) to disappear. This will make the demand system to be a function only of the quadratic expenditure term.

where the error term ε_i has been added to the right-hand side of (14) for estimation purposes. As is usual in demand system estimations, the error terms $\boldsymbol{\varepsilon} \equiv [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_K]$ are assumed to have a multivariate normal distribution with covariance matrix $\boldsymbol{\Sigma}$. We want to test the null hypothesis that the vector of coefficients $\boldsymbol{\theta}$ in $\left(\prod_{i=1}^{K} p_i^{\theta_i}\right)^{-1}$ is identically zero (i.e., H$_o$: $\boldsymbol{\theta} = \mathbf{0}$). The restricted model (with $\boldsymbol{\theta} = \mathbf{0}$) is easier to estimate than the unrestricted model, which makes the Lagrange Multiplier (LM) test an attractive approach.

Consider maximization of the log-likelihood subject to a set of constraints $c(\boldsymbol{\theta}) - r = 0$. Let κ be the Lagrange multiplier, and define the Lagrangean function:

$$\Lambda \quad = \ln L + \kappa(c(\boldsymbol{\theta}) - r) \tag{15}$$

where ln L is the log-likelihood function for commodity i given by:

$$\ln L \quad = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\sigma^2 - \frac{1}{2}\sum_{i=1}^{n}\left[\frac{(w_i - E(w_i))^2}{\sigma^2}\right]. \tag{16}$$

The first derivative of the log-likelihood function with respect to θ_i is:

$$\frac{\partial \ln L}{\partial \theta_i} \quad = \lambda_i\left(\prod_{i=1}^{K} p_i^{\theta_i}\right)^{-2}\left(\ln p_i \cdot p_i^{\theta_i} \cdot \prod_{j \neq i} p_j^{\theta_j}\right)\left\{\ln\left[\frac{x}{a(\mathbf{p})}\right]\right\}^2. \tag{17}$$

Evaluated at the null, H$_0$: $\boldsymbol{\theta} = \mathbf{0}$, the first derivative (17) becomes:

$$\frac{\partial \ln L}{\partial \theta_i} = \lambda_i \ln p_i \left\{\ln\left[\frac{x}{a(\mathbf{p})}\right]\right\}^2. \tag{18}$$

Based on equation (18), a test for statistical significance of prices in $b(\mathbf{p})$ reduces to adding price times expenditure-squared interaction terms ($\sum_{i=1}^{K} \ln p_i \cdot \{\ln[x / a(\mathbf{p})]\}^2$) to the demand model linear in expenditure (i.e., equation (5) with $\lambda_{\text{i}} = 0$) — the unrestricted model — and comparing it with the restricted model, which is just the QUAIDS expenditure share equations (5). To carry out this test, we first estimate the restricted model and obtain the residuals. These residuals are then regressed on all variables, including the price times expenditure-squared interaction terms. The R-squared, R_u^2, from this regression is used to compute the LM statistic, $LM = N \cdot R_u^2$. This *LM* statistic follows a Chi-squared distribution with the degrees of freedom equal to the number of restrictions being tested. For testing purposes, the translog price aggregator, $a(\mathbf{p})$, is approximated by the modified Stone price index suggested by Moschini (1995).

The LM test just discussed is useful for preliminary analysis of the data to determine whether the demand model should be specified with a quadratic (QUAIDS) or a linear (AIDS) expenditure variable. An obvious alternative would be to estimate the QUAIDS model and test for the statistical significance of the quadratic expenditure term.[3] However, the QUAIDS model is highly nonlinear and difficult to estimate. This LM test is a useful way to test parametrically whether or not the quadratic expenditure is necessary without having to estimate the highly nonlinear QUAIDS model.

2.2 Expenditure endogeneity

Most empirical demand analyses do not cover all products and services that households purchase. Data limitations, finite computer memory, and the increased complexity and time required for estimating large models make it necessary to abstract from a completely specified demand system containing a different equation for each of the myriad goods available in the market (LaFrance, 1991). The practice is typically to

---

[3] Also, one can use nonparametric methods to analyze the shape of the Engel curves (i.e., relationships between a commodity's budget share and total expenditure), as did Banks *et al*. (1997).

assume that preferences are separable and estimate a set of conditional demands for the goods of interest as functions of prices and total expenditure on these goods (Pollak, 1969). However, such a practice raises questions regarding the possibility of simultaneity bias in the budget share equations of the demand model. Total expenditure may be determined jointly with the expenditure shares of the individual commodities being analyzed, making it endogenous in the expenditure share equations. Also, expenditure endogeneity issues may arise whenever the household expenditure allocation process is correlated with other unobserved behavior not captured by the explanatory variables in the budget share equations, because these unobserved effects would be bundled in the error term. Estimation ignoring expenditure endogeneity may lead to inconsistent demand parameter estimates.

In cross-sectional demand studies, the common procedure to control for expenditure endogeneity is to use instrumental variables. With panel data, a number of possibilities to correct for unobserved heterogeneity are available, including linear transformations of the original model, such as through fixed effects and first differencing, to remove the unobserved heterogeneity component of the error term. However, such transformations are difficult to implement with nonlinear models. With demand models such as QUAIDS that are derived from consumer utility maximization theory, nonlinearities are inevitable making it difficult to implement these linear transformations. In this study, we follow Bundell and Robin (1999) and control for endogeneity using an extension of the limited information augmented regression technique suggested by Hausman (1978).

To illustrate how the augmented regression technique works, consider the regression of y_1, the dependent variable, on a set of exogenous explanatory variables, $\mathbf{z}$, and an endogenous explanatory variable, y_2, i.e., $y_1 = \mathbf{z}'\rho + \pi y_2$.[4] Also, suppose an instrumental variable, z_2, exists for y_2. Correction for the endogeneity of y_2 using the control function approach proceeds in two steps. The first step involves estimating a reduced form regression of the endogenous variable on a set of instrumental variables, where the set of instrumental variables include all the other exogenous explanatory variables (i.e., regress y_2 on $\mathbf{z}$ and z_2). The residuals, $\hat{v}$, from this first-stage regression are then included as an additional explanatory variable in the original y_1 equation. The OLS estimator of the parameters ρ and π in this augmented regression is identical to the Two-Stage Least Squares (2SLS) estimator (Blundell and Robin, 1999). Moreover, testing for the significance of the coefficient on $\hat{v}$ is a test for the exogeneity of y_2. Following Banks *et al*. (1997), we use total household income and its square as instruments for expenditure (and expenditure squared).

3. Data

Data used in this study comes from the KwaZulu-Natal Income Dynamics Study (KIDS). KIDS is a panel dataset comprising three surveys: the 1993 Project for Statistics on Living Standards and Development (PSLSD) survey, and the 1998 and 2004 surveys which interviewed households from the 1993 PSLSD survey who reside in KwaZulu-Natal Province.

---

[4] For illustration purposes, we consider the case of *one* endogenous variable and *one* instrumental variable. The case of multiple endogenous variables and multiple instruments can be handled in a straightforward way using the basic framework explained here.

The PSLSD is a nation-wide survey undertaken in the last half of 1993 by a consortium of South African survey groups and universities, including the South African Labor and Development Research Unit of the University of Cape Town and the World Bank. The main instrument was a comprehensive household survey collecting data on a broad array of socio-economic conditions of households, including their food and non-food expenditures. Respondents were selected using a two-stage, self-weighting design. In the first stage, clusters were chosen proportional to size from census enumerator districts; the census enumerator districts were based on the 1991 population census. In the second stage, all households in each chosen cluster were enumerated and then a random sample of them selected. In addition to the household questionnaire, a community questionnaire was administered in each cluster to collect data on prices for a detailed list of food products commonly purchased by households. The recall period in all the three panel surveys is one month.

In 1998, households surveyed by the PSLSD in KwaZulu-Natal Province were reinterviewed by the KIDS survey. The 1998 KIDS survey was undertaken by a consortium including the University of Natal, the University of Wisconsin-Madison, and the International Food Policy Research Institute. The third KIDS survey was undertaken in 2004, this time with the 1998 consortium expanded to include the London School of Hygiene and Tropical Medicine (LSHTM) and the Norwegian Institute of Urban and Regional Studies (NIBR). Both the 1998 and 2004 surveys tracked and interviewed households who had moved. The 1998 target sample (i.e., the sample that would have resulted in the absence of attrition) from the KwaZulu Natal province was 1,354 households. Due to problems of sample attrition, and after removing observations deemed unusable for the current purpose (such as households reporting zero consumption on all food items, or those missing critical information, such as expenditure data), a total of 727 observations per survey are used. While the 1998 and 2004 KIDS interviewed new households, such as those that split from the original 1993 households, this study is restricted to the households that were interviewed in all the three surveys, resulting in a total sample of 2,181 (3×727) observations.[5]

KwaZulu-Natal is the most populous province in South Africa, constituting approximately 20% of South Africa's population. The economic, social, and racial stratification of KwaZulu-Natal mirrors that of the country as a whole: the province includes a wealthy metropolitan area, Durban, poor townships surrounding it and a poor and largely rural former homeland, KwaZulu. Also, poverty and inequality in the province are relatively similar to those at the national level (Woolard *et al*., 2002), so that results of the analysis should provide important insights about the conditions in other provinces.

Table 1 provides a summary of the food expenditure patterns of the sampled households, including the differences across income groups, and rural and urban areas.[6] Grains constitute the largest share of household total food expenditure, ranging from about 26 and 24% among the high income and urban households to 37% among the low income and rural households. The share of grains in the households' budgets is lower at higher income levels

---

[5] Tests for attrition (not reported here) indicated that attrition does not affect consistent estimation of the demand parameter estimates.

[6] To create the income groups, the (CPI-deflated) income of each household is averaged over the three panel years, and then households are ranked from lowest to highest based on their averaged incomes. The households are then divided into three income groups, with households in each income group comprising about one-third of the total sample.

Table 1. Average Expenditure Shares by Income Group and Region

	Entire sample	Income groups			Regions	
		Low	Middle	High	Rural	Urban
Grains	0.33	0.37	0.35	0.26	0.37	0.24
Meat, fish	0.23	0.21	0.22	0.26	0.21	0.27
Fruits, vegetables	0.18	0.18	0.18	0.16	0.18	0.17
Dairy products	0.07	0.05	0.06	0.08	0.05	0.09
Oils, butter, fats	0.05	0.05	0.05	0.05	0.05	0.05
Sugar, sugar products	0.05	0.06	0.05	0.03	0.05	0.04
Other food	0.10	0.07	0.08	0.14	0.08	0.14
Total household food expenditure	807.29	584.97	746.51	1089.90	725.94	967.34
Total household income	2666.48	774.34	1696.70	6064.22	1981.55	4553.35
Food Expenditure as % of income	0.30	0.75	0.44	0.18	0.37	0.21
Sample size	2181	729	723	729	1446	735

Table 2. Summary Statistics of Household Characteristics by Income Group and Region

Variable	Entire sample	Income Groups			Regions	
		Low	Middle	High	Rural	Urban
Household size	6.79	6.37	7.41	6.61	7.34	5.72
Education of head	5	3.15	3.89	7.83	3.76	7.31
Age of household head	54.67	53.68	56.96	53.39	55.58	52.86
Proportion male headed	0.59	0.51	0.56	0.71	0.57	0.64
Proportion rural	0.66	0.88	0.74	0.36	-	-
Proportion black (race)	0.85	0.97	0.94	0.65	1.00	0.57

The budget share of meat products, a more expensive source of calories, is higher among the high income and relatively affluent urban households. The mean monthly income of high income households is seven times more than that of low income households, reflecting the generally high wealth inequality in South Africa.

Summary statistics of household demographic characteristics are presented in Table 2. On average, urban households are of smaller size, headed by younger males with high levels of education. Most of these characteristics are shared by high-income households, except the latter have larger family sizes. The rural and low-income groups comprise mainly black households.

4. Empirical estimation

4.1. LM test results for quadratic expenditure specification

We conduct the LM tests for AIDS versus QUAIDS first in the individual budget share equations for each commodity, and then across all equations estimated jointly as a system. The first column of Table 3 reports the results of these tests. The null hypothesis that the coefficient on the quadratic expenditure term is independent of prices is rejected (at the 5% significance level) in all individual budget share equations, except that for meat and fish. The null hypothesis that this coefficient is independent of prices across all the budget share equations is rejected, based on the χ^2 test (=175.56) from the SUR estimation. Based on these results, we conclude that the rank 3 QUAIDS specification is the preferred over AIDS for modeling food demand in this study.

As explained in section 2.2, total expenditure may be endogenous in the budget share equations, and this may affect the results of the LM tests. Using log total household income ($\ln m$) and log total household income squared $(\ln m)^2$ as instrumental variables (IVs) for log total household food expenditure ($\ln x$) and log total household food expenditure squared $(\ln x)^2$, we explicitly test for the endogeneity of $\ln x$ in the budget share equations.

Table 3. Tests for Nonlinearity of the Demand System based on Statistical Significance of Prices of the Coefficient on the Quadratic Expenditure Term

Equation-by-equation tests

Commodity	**OLS**	**IV-2SLS (p-value)**
Grains	30.31 (0.0001)	26.17 (0.0005)
Meat and fish	6.32 (0.5024)	33.58 (0.0000)
Fruits and vegetables	21.81 (0.0027)	21.59 (0.0030)
Dairy	21.37 (0.0032)	9.59 (0.2126)
Oils, butter, and other fats	14.61 (0.0413)	11.34 (0.1244)
Sugar	21.81 (0.0027)	12.43 (0.0872)
Other	85.93 (0.0000)	26.83 (0.0003)

Equation-system tests (i.e., across all budget share equations)

	SUR	**3SLS**
χ^2 (p-value)	175.56 (0.0000)	55.36 (0.0811)

A good IV must meet two standard conditions: the *relevanc*e condition, which requires that it must be sufficiently correlated with the endogenous variable, and the *exogeneity* condition, which requires that it must not be correlated with the error term in the demand model. The former condition is testable, and the latter cannot be tested. To test for the relevance condition, we estimate two reduced form regressions, one for $\ln x$ and the other for $(\ln x)^2$, since both these variables enter the QUAIDS model. Results of these reduced form regressions are reported in the last row of Table I in the appendix.

Based on simple t-tests, the individual coefficients on ln m and (ln m)^2 are significantly different from zero in both reduced form regressions, providing some evidence that income can be a good instrument for expenditure. A formal test for the relevance of ln m and (ln m)^2 as IVs for ln x and (ln x)^2 involves testing for the joint significance of the coefficients on both ln m and (ln m)^2 in the reduced form regressions. The results of these tests are reported in Table II in the appendix. Based on these results, it can be concluded that ln m and (ln m)^2 are relevant instruments for ln x and (ln x)^2, and hence, the former will be used as IVs for the latter in the analyses that follow (the exogeneity assumption is, of course, maintained). The residual-based procedure is used to test for the endogeneity of expenditure in the budget share equations (see Wooldridge (2002): 118-122). The results of these tests are reported in Table II in the appendix. The null hypothesis of expenditure exogeneity is rejected in all individual budget share equations, except those for fruits and vegetables and oils and fats. The null hypothesis that expenditure is endogeneous across all budget share equations is rejected ($\chi^2 = 70.07$, p = 0.0000). To adjust for expenditure endogeneity, we use instrumental variables two-state least squares (IV-2SLS) to estimate the budget share equations, and conduct the LM tests for quadratic expenditure specification as before. Results of these tests are reported in the second column of Table 3.

Based on the results of the LM tests from IV-2SLS regressions, the null hypothesis that the budget share equation for meat and fish is linear expenditure in expenditure is rejected. We consider these results (based on IV-2SLS) to be more reliable, given the finding from the endogeneity tests (Table II) that expenditure is endogenous in the budget share equation for meat and fish. Contrary to the LM test results based on the equation-by-equation OLS estimations, the null hypothesis that expenditure is linear in the dairy equation is not rejected. Consistent with the SUR estimations above, the null hypothesis that the coefficient on the quadratic expenditure term is independent of prices across all budget share equations is also rejected in the three-stage least squares estimations (adjusting for expenditure endogeneity). So, the tests in both the endogeneity-adjusted and endogeneity-unadjusted estimations provide evidence in favor of the QUAIDS model specification. Hence, all the demand analyses that follow will be based on the endogeneity-adjusted QUAIDS model.

4.2. Demand model results

The QUAIDS model is estimated using pooled maximum likelihood (ML), with theoretical restrictions of adding-up, homogeneity, and symmetry imposed during estimation.[7] We first test conduct an additional test for the AIDS versus QUAIDS specification based on the statistical significance of lambda (λ_i) in the coefficient on the quadratic expenditure term, $\lambda_i / b(\mathbf{p})$. This test is based on the fact that QUAIDS nests AIDS, so that once QUAIDS has been estimated, a test for whether or not AIDS is the appropriate model specification involves simply checking for the statistical significance of the quadratic expenditure term. Results of the tests for the significance of λ_i are reported in Table 4, both for the entire sample (first column) and for the rural-urban and income groups sub-samples. These tests are conducted in the endogeneity-unadjusted QUAIDS model. The reason for conducting the tests using the endogeneity-unadjusted model, as opposed to the model with endogeneity adjusted for, is to avoid problems of

---

[7] A more general approach would be to estimate this system allowing cross periods correlations, such as using the generalized method of moments (GMM) estimator.

inferential invalidity caused by generated regressors (Wooldridge, 2002; pp. 115-118).[8] So, the results of these tests are comparable with those from OLS estimations in Table 3.

In the pooled data (i.e., pooled across all households—rural, urban, and income groups), the null hypothesis that λ_i is zero is rejected in the budget share equations of five of the seven food groups. The hypothesis that λ is zero across all budget share equations is rejected ($\chi^2 = 72.03$, p = 0.000). The null hypothesis that λ is zero is not rejected in the budget share equations for meat and fish and oils and fats. The finding that λ is not significant in the budget share equation for meat and fish is consistent with the results of LM tests in Table 3, which indicated that this budget share equation does not need a quadratic expenditure term. The results in the oils and fats equation differ from LM-based tests results, because the latter provided statistical evidence favoring the inclusion of the quadratic expenditure term. However, the approaches followed in constructing the two tests differ, so their leading to different conclusions regarding which budget share equations are linear in expenditure and which are not is not necessarily unexpected. Another factor that may lead to result differences is the fact that ln $a(\mathbf{p})$ is approximated by a linear price index in the regressions upon which the LM tests are based. In the case of the system wide tests, both the LM and likelihood ratio approaches lead to a consistent conclusion that the system of demand equations should be estimated using QUAIDS, not AIDS.

Consistent with the findings in the pooled data, equation-by-equation tests for statistical significance of lambda in the rural-urban and income groups sub-samples

---

[8] This problem arises here because the use of the augmented regression approach involves including the residuals from the reduced form regressions as regressors in the demand model. But because these residuals are generated using the same data used for demand estimation, their inclusion as regressors invalidates standard errors and test statistics on other regressors (parameters can still be estimated consistently).

Table 4. Tests for Quadratic Expenditure Specification and Endogeneity

Statistical Significance of Lambda: χ^2 tests[1,2]

Commodity	**Entire sample**	**Rural**	**Urban**	**Low**	**Middle**	**High**
Grains	17.55	23.07	5.22	0.74	7.65	0.01
	(0.0000)	(0.0000)	(0.0223)	(0.3891)	(0.0057)	(0.9296)
Meat, fish	0.45	8.07	35.08	2.80	0.02	6.01
	(0.5017)	(0.0045)	(0.0000)	(0.0943)	(0.9246)	(0.0142)
Fruits, vegetables	3.76	2.86	1.68	0.16	1.40	3.11
	(0.0525)	(0.0906)	(0.1953)	(0.6929)	(0.2367)	(0.0779)
Dairy	7.83	2.56	8.55	4.01	0.01	11.74
	(0.0051)	(0.1098)	(0.0035)	(0.0453)	(0.9389)	(0.0006)
Oils, butter, fats	0.95	0.11	4.87	1.59	4.12	0.24
	(0.3309)	(0.7366)	(0.0273)	(0.2078)	(0.0424)	(0.6211)
Sugar	8.77	8.36	2.95	5.78	5.40	2.23
	(0.0031)	(0.0038)	(0.0858)	(0.0162)	(0.0202)	(0.1352)
Other	48.84	12.59	24.08	2.21	13.85	7.34
	(0.0000)	(0.0004)	(0.0000)	(0.1368)	(0.0002)	0.0067
All equations	72.03	*41.40*	*68.30*	*14.22*	*24.16*	*29.73*
	(0.0000)	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*	*(0.0000)*

F-tests for the Relevance of Instruments

Expenditure	103.96	48.92	59.47	18.06	14.11	16.56
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Expenditure-squared	104.66	46.07	60.80	16.65	13.59	16.45
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)

Expenditure Endogeneity: χ^2-tests[1,3]

Grains	11.78	1.28	11.48	15.72	6.18	0.19
	(0.0028)	(0.5275)	(0.0000)	(0.0004)	(0.0455)	(0.9076)
Meat, fish	3.33	7.02	0.73	0.30	0.85	9.61
	(0.1892)	(0.0299)	(0.6926)	(0.8601)	(0.6547)	(0.0082)
Fruits, vegetables	3.81	1.77	7.91	1.09	4.96	10.21
	(0.1488)	(0.4133)	(0.0191)	(0.5807)	(0.0839)	(0.0061)
Dairy	0.15	1.89	3.91	1.39	1.49	1.18
	(0.9292)	(0.3878)	(0.1415)	(0.5001)	(0.4736)	(0.5547)
Oils, butter, fats	8.96	1.00	11.05	7.33	3.61	1.18
	(0.0113)	(0.6072)	(0.0040)	(0.0257)	(0.1645)	(0.5547))
Sugar	4.09	4.19	9.16	10.39	0.72	0.89
	(0.1295)	(0.1230)	(0.0103)	(0.0055)	(0.6962)	(0.6399)
Other	18.38	3.00	18.98	30.55	8.02	1.43
	(0.0001)	(0.2233)	(0.0001)	(0.0000)	(0.0181)	(0.4901)
All equations	42.07	*17.72*	*54.61*	*57.43*	*22.15*	*20.34*
	(0.0000)	*(0.1245)*	*(0.0000)*	*(0.0000)*	*(0.0359)*	*(0.0609)*

[1] p-value (prob > χ^2) in parentheses

[2] Based on QUAIDS estimation without adjustment for expenditure endogeneuty

[3] Based on QUAIDS estimation with reduced form residuals augmented to each of the budget shares equations

provides evidence in favor of QUAIDS in some budget share equations and AIDS in others. However, when all the budget share equations are considered jointly as a system, the evidence in favor of QUAIDS is robust across all the five sub-samples. The strength of the evidence in support of QUAIDS differs across the sub-samples in an interesting way. There is a clear rural-urban difference, with the statistical evidence in support of QUAIDS stronger in the urban than the rural sample. The statistical evidence in favor of QUAIDS also tends to be weaker in the individual income groups samples than in the pooled data. A possible explanation is that because the role of the quadratic expenditure term is to capture the variability in expenditure elasticity along the income spectrum of households, by grouping households with similar income groups together (i.e., making the sample more income homogenous) reduces the need to include a quadratic expenditure term.

Table 4 also reports the results of the tests for expenditure ($\ln x$) endogeneity in each of the rural, urban, and income-groups sub-samples. The results of the tests for the relevance of $\ln m$ (income) and $(\ln m)^2$ as instruments for $\ln x$ and $(\ln x)^2$ indicate that $\ln m$ explains a significant portion of the variation in $\ln x$, so that assuming the exogeneity condition is satisfied, $\ln m$ can be used a valid instrument for $\ln x$. Based on the results of the χ^2 tests (reported at the bottom part of Table 4), the null hypothesis of expenditure endogeneity is rejected in all the sub-samples except the rural, in which statistical evidence against expenditure exogeneity is weaker ($p = 0.1245$).

Based on the model specification and endogeneity tests results, the analyses that follow (i.e., of rural, urban, and income groups demands) will be based on the endogeneity-adjusted QUAIDS model. The demographic variables included in the

QUAIDS model are: household size, race, education, age, and gender of household head, as well as the month of the survey. A rural-urban dummy is included when pooled data are used. For reasons of space, we do not report all the ML parameter estimates of the demand model. Table III in the appendix reports coefficients on the price and expenditure variables. In the pooled data, 19 of the 28 price effects are significantly different from zero at the 10 percent significance level, suggesting that there are considerable quantity changes in response to movements in relative prices. Although not explicitly shown in the table, in both the pooled data and the five sub-samples, larger-sized households consume more grains and less meat and fish and fruits and vegetables. The month of survey is significant (at the 10 percent level) across all the model estimations, indicating the importance of seasonality in food purchase and consumption patterns.

Our discussion will focus on elasticities, because price and income effects are better discussed in terms of price and expenditure elasticities. Table 5 reports the estimated expenditure and own-price elasticities. The elasticities are computed at the sample means.

All the estimated elasticities conform to a priori expectations— expenditure elasticities are all positive and own-price elasticities negative. Thus, based on signs on the expenditure elasticities, all seven food commodities are normal, so that their demand increases as total household expenditure increases. In the pooled data, meat and fish and dairy are luxuries, with expenditure elasticities in excess of unity. Meat and fish and dairy are also more price elastic than all the other food groups. This indicates the degree of latitude that households have in responding to the changes in the prices of these foods. However, when only the substitution effects are considered, meat becomes less price elastic, as shown by the inelastic compensated own-price elasticity.

Table 5. Estimated Expenditure and Own-Price Elasticities

Expenditure elasticities[1]

Commodity	Entire sample	Rural	Rural[2] No endog. adj.	Urban	Low	Middle	High
Grains	0.9114	1.0358	1.1197	0.4981	1.3878	1.1088	0.6770
Meat, fish	1.2336	1.0030	1.0341	1.4517	0.7654	1.2302	1.1752
Fruits, vegetables	0.8249	0.8355	0.8222	0.7865	0.7450	0.7848	0.8961
Dairy	1.3622	1.4537	1.3046	1.3735	2.0498	1.0361	1.6023
Oils, butter, fats	0.8367	0.7215	0.7871	0.8989	0.9997	0.8998	0.6915
Sugar	0.8114	0.8195	1.0324	0.6130	1.9485	0.9376	0.2091
Other	0.9982	1.1965	0.6617	1.1703	1.0481	0.4645	1.3631

Marshallian/uncompensated own-price elasticities

Commodity	Entire sample	Rural	Rural No endog. adj.	Urban	Low	Middle	High
Grains	-0.9550	-1.0013	-1.0413	-0.8904	-1.2028	-1.0059	-0.9044
Meat, fish	-1.0725	-1.0798	-1.0839	-1.1680	-0.9934	-1.0962	-1.0846
Fruits, vegetables	-0.9379	-0.9625	-0.9613	-0.8783	-0.9415	-0.9274	-0.9703
Dairy	-1.1659	-1.2279	-1.1680	-0.9931	-1.2190	-1.0994	-1.1557
Oils, butter, fats	-1.0306	-1.0511	-1.0669	-1.0578	-0.9142	-1.2298	-0.9230
Sugar	-0.9059	-0.9332	-0.9255	-0.9366	-0.9838	-0.9270	-0.8655
Other	-0.7353	-0.7887	-0.8849	-0.9644	-0.8852	-0.7186	-0.6824

Hicksian/compensated own-price elasticities

Commodity	Entire sample	Rural	Rural No endog. adj.	Urban	Low	Middle	High
Grains	-0.6582	-0.6214	-0.6305	-0.7686	-0.5949	-0.6201	-0.7264
Meat, fish	-0.7878	-0.8681	-0.8656	-0.7769	-0.8343	-0.8240	-0.7755
Fruits, vegetables	-0.7925	-0.8115	-0.8126	-0.7469	-0.8040	-0.7861	-0.8234
Dairy	-1.0763	-1.1473	-1.0956	-0.8749	-1.1121	-1.0332	-1.0253
Oils, butter, fats	-0.9860	-1.0131	-1.0255	-1.0086	-0.8587	-1.1829	-0.8868
Sugar	-0.8664	-0.8882	-0.8688	-0.9142	-0.8713	-0.8774	-0.8581
Other	-0.6359	-0.6951	-0.8331	-0.7986	-0.9648	-0.6805	-0.4905

[1] All elasticities are computed from endogeneity-adjusted QUAIDS model

[2] These are the expenditure and own-price elasticities computed from endogeneity-unadjusted QUAIDS model.

Dairy remains price elastic when both the uncompensated and the compensated elasticity estimates are considered. The differences in the estimated expenditure elasticities between rural and urban samples is quite substantial. For urban households, a 1% increase in total food expenditure leads to only about half a percentage increase in the budget share of grains. It is very different with the rural households, where the same 1%

expenditure increase leads to about 1.12% increase in expenditure on grains. This is one of the reasons why it is necessary to examine expenditure patterns of rural households separately from urban households.

The magnitudes of differences in the expenditure elasticities are also large across income groups. Expenditure elasticities for grains range from 0.6770 for high income households to 1.3878 for low income households, while for sugar and sugar products they range from 1.9485 for low income households to only 0.2091 for high income households. This reaffirms the need for disaggregated analysis of food expenditure patterns by income groups in a country with high income inequalities like South Africa. Nevertheless, the estimated expenditure elasticities are as one would expect a priori. Expenditure elasticity of grains is highest among low income households, but lower among the middle and high income households.

4.3 Demand estimates with correction for zero expenditures

In the pooled data, the problem of zero expenditures is severe for the dairy commodity. About 14% of the dairy budget shares are zeros. Apart from dairy, the percentages of observations with zero expenditures is very low (4% at most) in the pooled data. Non-purchase of dairy products is higher among rural households (18%) and among households in the lower income brackets (24% for low-income households and 12% for middle-income households). We adjust for zero expenditures in the dairy commodity using a two-step procedure proposed by Shonkwiler and Yen (1999). In the first step, a single equation probit model is estimated in order to compute the probability and the cumulative density values. In the second step, the demand system is estimated with the budget shares of the dairy commodity weighted by the cumulative density values, and probability density values included as an additional regressor in the budget share equation for the dairy commodity. The discussion here focuses on the statistical significance of the coefficient on the probability density values (δ), and on the impact on the elasticity estimates for the dairy commodity of controlling for zero expenditures. The results of these tests are reported in the last row of Table 6.[9]

Table 6. Expenditure and Own-Price Elasticities for Dairy Adjusting for the zero-expenditure problem

	Household Group			
	Rural	**Low**	**Middle**	**All**
Expenditure elasticity	1.1918	1.6941	0.8530	1.1373
Marshallian own-price elasticity	-1.1012	-1.1662	-0.9953	-1.0320
Hicksian own-price elasticity	-1.0351	-1.0779	-0.9408	-0.9572
t-ratio (δ=0 vs. $\delta \neq 0$) (p-value)	-6.02 (0.000)	-3.19 (0.001)	-4.75 (0.000)	-7.15 (0.000)

As the results show, δ is significant in the budget share for dairy in the pooled sample and in each of the three subsamples, indicating that the additional information provided by the probability density values explains a significant part of the variation in the budget share of dairy. Comparing the elasticity estimates in Table 6 with those in Table 5, it can be seen that the correction for the zero expenditures has the effect of decreasing the magnitudes of the expenditure elasticity estimates for dairy. The reduction in the estimated expenditure elasticities is quite large in all cases, and large enough to change

---

[9] As before, we avoid problems of inferential invalidity associated with generate regressors, (see Wooldridge (2002), pp. 115-118) by testing for the statistical significance of δ in the endogeneity-unadjusted demand model.

the classification of dairy from luxury to necessity in the case of middle income households.

5. Summary and Conclusions

This paper analyzed food expenditure patterns in South Africa, taking into account differences in preferences across rural and urban households, as well as across income groups. The LM tests developed in this paper provided evidence in favor of the QUAIDS model, and hence, the QUAIDS model was used to estimate the demand functions for seven food aggregates. Expenditure was found to be endogenous, and adjustment for this endogeneity was done using the augmented regression approach. The problem of zero expenditures was controlled for using a two-step procedure proposed by Shonkwiler and Yen (1999). In the pooled data, meat and fish and dairy were found to be luxuries, with expenditure elasticities in excess of unity, while the other commodities were found to be necessities. Demand patterns between rural and urban households, as well as across income groups, differ substantially. For urban households, a 1% increase in total food expenditure leads to only about half a percentage increase in the budget share of grains. It is very different with the rural households, where the same 1% expenditure increase leads to about 1.12% increase in expenditure on grains. Expenditure elasticities for grains range from 0.6770 for high income households to 1.3878 for low income households, while for sugar and sugar products they range from 1.9485 for low income households to only 0.2091 for high income households. These results show that that in a country like South Africa with high income inequality and huge divide in living standards between rural and urban households, it is necessary to examine consumption patterns at levels of disaggregation similar to the one followed in this study, and to account for quadratic expenditure effects, expenditure endogeneity, and zero expenditures in order to get reliable and consistent results.

APPENDIX: ADDITIONAL RESULTS

Table I. The Estimated Reduced Forms for log Total Household Food Expenditure (ln *x*) and log Total Hhousehold Food Expenditure-Squared ((ln *x*)^2)

	Dependent variable	
Variable	ln *x* (std. err)	(ln *x*)^2 (std. err)
Constant	6.8458 (0.2795)	47.7117 (3.2928)
Price of grains	-0.0450 (0.0608)	-0.4593 (0.7164)
Price of meat & fish	-0.0795 (0.0503)	-0.7704 (0.5927)
Price of fruits & vegetables	-0.1229 (0.0405)	-0.1397 (0.4772)
Price of dairy	-0.0582 (0.0369)	-0.5612 (0.4357)
Price of oils, butter & fats	0.0593 (0.0450)	0.5923 (0.5889)
Price of sugar	-0.3028 (0.0411)	-3.6918 (0.4849)
Price of other foods	-0.2874 (0.0511)	-3.3802 (0.6022)
Total household income	-0.0580 (0.0635)	-1.1041 (0.7491)
Total household income^2	0.0172 (0.0048)	0.2349 (0.0568)
Household size	0.0455 (0.0031)	0.5485 (0.0365)
Race	-0.3862 (0.0420)	-4.8390 (0.4944)
Rural	0.0086 (0.0307)	0.1909 (0.3520)
Education	0.0126 (0.0027)	0.1461 (0.0317)
Age	0.0020 (0.0009)	0.0244 (0.0104)
Gender	0.0858 (0.0235)	1.0436 (0.2765)
Survey month	-0.0026 (0.0074)	-0.0395 (0.0867)
R^2	*0.3694*	*0.3751*
Test for the relevance of income and income^2 as instruments for expenditure and expenditure^2		
F stat. (p-value)	103.96 (0.0000)	104.66 (0.0000)

Table II. Results of the Test for the Endogeneity of log Total Household Expenditure

Equation-by-equation tests

Commodity	**F-tests (p-value)**
Grains	7.94 (0.0004)
Meat and fish	14.62 (0.0000)
Fruits and vegetables	1.36 (0.2571)
Dairy	4.59 (0.0102)
Oils, butter, and other fats	0.85 (0.4275)
Sugar	4.40 (0.0124)
Other	10.58 (0.0000)

Equation-system tests (i.e., across all budget share equations)

	SUR
χ^2 (p-value)	70.07 (0.0000)

Table III. Price and Expenditure Parameter Estimates of the QUAIDS model[1]

Variable	**Eq.**[2]	**Entire sample**	**Rural**	**Urban**	**Low**	**Middle**	**High**
Constant	α_1	0.0604	0.1403	0.0542	0.2029	0.1201	0.0614
		(0.0280)	(0.0413)	(0.0366)	(0.0670)	(0.0534)	(0.0388)
	α_2	0.3886	0.3673	0.3779	0.3556	0.3845	0.3490
		(0.0262)	(0.0364)	(0.0393)	(0.0587)	(0.0493)	(0.0416)
	α_3	0.1805	0.2251	0.1555	0.1597	0.2088	0.1978
		(0.0188)	(0.0270)	(0.0259)	(0.0441)	(0.0356)	(0.0265)
	α_4	0.0680	0.0371	0.0606	0.0573	0.0414	0.0830
		(0.0151)	(0.0211)	(0.0233)	(0.0296)	(0.0332)	(0.0224)
	α_5	0.0365	0.0489	0.0223	0.0441	0.0316	0.0172
		(0.0091)	(0.0126)	(0.0143)	(0.0205)	(0.0129)	(0.0176)
	α_6	0.0008	0.0272	-0.0192	0.0263	-0.0069	-0.0094
		(0.0075)	(0.0113)	(0.0093)	(0.0206)	(0.0131)	(0.0085)
	α_7	0.2651	0.1540	0.3486	0.1539	0.2205	0.3011
		(0.0237)	(0.0298)	(0.0418)	(0.0499)	(0.0411)	(0.0411)
Expenditure	β_1	-0.1254	-0.1228	-0.0656	-0.0619	-0.1126	-0.0932
		(0.0155)	(0.0205)	(0.0232)	(0.0355)	(0.0306)	(0.0250)
	β_2	0.0329	0.0692	-0.0449	0.0579	0.0544	-0.0606
		(0.0145)	(0.0181)	(0.0242)	(0.0310)	(0.0291)	(0.0267)
	β_3	0.0046	0.0183	-0.0178	-0.0146	0.0144	0.0417
		(0.0101)	(0.0127)	(0.0169)	(0.0222)	(0.0200)	(0.0168)
	β_4	-0.0036	-0.0112	-0.0115	-0.0167	-0.0014	-0.0088
		(0.0084)	(0.0106)	(0.0147)	(0.0159)	(0.0197)	(0.0150)
	β_5	-0.0095	-0.0006	-0.0271	-0.0068	-0.0206	-0.0063
		(0.0049)	(0.0062)	(0.0085)	(0.0109)	(0.0077)	(0.0112)
	β_6	-0.0235	-0.0211	-0.0302	-0.0173	-0.0273	-0.0196
		(0.0041)	(0.0057)	(0.0054)	(0.0109)	(0.0077)	(0.0053)
	β_7	0.1245	0.0682	0.1973	0.0594	0.0932	0.1468
		(0.0132)	(0.0150)	(0.0256)	(0.0266)	(0.0247)	(0.0264)
Prices	γ_{11}	0.0216	0.0046	0.0203	0.0001	0.0132	0.0221
		(0.0135)	(0.0180)	(0.0173)	(0.0246)	(0.0220)	(0.0188)
	γ_{21}	-0.0134	-0.0059	-0.0093	-0.0035	0.0002	-0.0025
		(0.0097)	(0.0129)	(0.0124)	(0.0183)	(0.0167)	(0.0137)
	γ_{31}	0.0187	0.0137	0.0187	0.0110	0.0222	0.0130
		(0.0050)	(0.0063)	(0.0072)	(0.0094)	(0.0083)	(0.0072)
	γ_{41}	0.0122	0.0154	0.0028	0.0162	0.0078	0.0045
		(0.0051)	(0.0066)	(0.0075)	(0.0087)	(0.0092)	(0.0078)
	γ_{51}	0.0047	0.0018	0.0123	0.0012	-0.0009	0.0072
		(0.0040)	(0.0053)	(0.0063)	(0.0076)	(0.0054)	(0.0073)
	γ_{61}	-0.0073	-0.0104	0.0020	-0.0187	-0.0093	0.0011
		(0.0032)	(0.0044)	(0.0041)	(0.0069)	(0.0051)	(0.0037)
	γ_{71}	-0.0364	-0.0192	-0.0468	-0.0063	-0.0332	-0.0453
		(0.0093)	(0.0112)	(0.0141)	(0.0160)	(0.0147)	(0.0146)
	γ_{22}	0.0043	-0.0121	0.0062	-0.0116	-0.0014	-0.0030
		(0.0105)	(0.0133)	(0.0141)	(0.0110)	(0.0182)	(0.0163)
	γ_{32}	-0.0068	-0.0026	-0.0099	-0.0031	-0.0032	-0.0122
		(0.0045)	(0.0056)	(0.0068)	(0.0087)	(0.0078)	(0.0067)
	γ_{42}	0.0023	-0.0056	0.0030	-0.0055	-0.0082	0.0204
		(0.0044)	(0.0056)	(0.0066)	(0.0077)	(0.0082)	(0.0072)
	γ_{52}	-0.0022	0.0037	-0.0137	-0.0044	0.0026	-0.0026
		(0.0040)	(0.0048)	(0.0065)	(0.0072)	(0.0054)	(0.0080)

Table III. (Continued)

	γ_{62}	0.0013	0.0064	-0.0043	0.0136	0.0014	-0.0056
		(0.0029)	(0.0038)	(0.0040)	(0.0062)	(0.0047)	(0.0038)
	γ_{72}	0.0146	0.0161	0.0279	0.0144	0.0086	0.0056
		(0.0077)	(0.0089)	(0.0128)	(0.0136)	(0.0125)	(0.0133)
	γ_{33}	0.0058	0.0007	0.0150	0.0032	0.0061	0.0028
		(0.0037)	(0.0045)	(0.0062)	(0.0076)	(0.0064)	(0.0054)
	γ_{43}	-0.0079	-0.0008	-0.0111	-0.0055	-0.0101	-0.0061
		(0.0028)	(0.0034)	(0.0049)	(0.0048)	(0.0052)	(0.0046)
	γ_{53}	0.0044	0.0037	0.0059	0.0002	0.0040	0.0081
		(0.0022)	(0.0026)	(0.0040)	(0.0042)	(0.0030)	(0.0042)
	γ_{63}	-0.0053	-0.0061	-0.0011	-0.0069	-0.0074	-0.0010
		(0.0020)	(0.0025)	(0.0032)	(0.0044)	(0.0032)	(0.0025)
	γ_{73}	-0.0089	-0.0017	-0.0174	0.0010	-0.0115	-0.0046
		(0.0047)	(0.0053)	(0.0086)	(0.0084)	(0.0075)	(0.0078)
	γ_{44}	-0.0094	-0.0086	0.0028	-0.0079	-0.0062	-0.0086
		(0.0037)	(0.0047)	(0.0062)	(0.0060)	(0.0075)	(0.0064)
	γ_{54}	0.0033	0.0036	0.0018	0.0058	0.0063	-0.0024
		(0.0021)	(0.0026)	(0.0038)	(0.0037)	(0.0029)	(0.0045)
	γ_{64}	0.0064	0.0079	0.0016	0.0049	0.0099	0.0017
		(0.0017)	(0.0023)	(0.0026)	(0.0035)	(0.0029)	(0.0024)
	γ_{74}	-0.0069	-0.0050	-0.0009	-0.0081	0.0004	-0.0095
		(0.0045)	(0.0053)	(0.0079)	(0.0075)	(0.0080)	(0.0078)
	γ_{55}	-0.0019	-0.0040	-0.0029	0.0048	-0.0119	0.0038
		(0.0036)	(0.0043)	(0.0074)	(0.0064)	(0.0047)	(0.0079)
	γ_{65}	-0.0045	-0.0059	-0.0037	-0.0047	-0.0070	0.0001
		(0.0020)	(0.0025)	(0.0039)	(0.0040)	(0.0028)	(0.0032)
	γ_{75}	-0.0037	-0.0029	0.0003	-0.0030	0.0068	-0.0142
		(0.0037)	(0.0044)	(0.0066)	(0.0068)	(0.0051)	(0.0073)
	γ_{66}	0.0047	0.0045	0.0029	0.0040	0.0043	0.0049
		(0.0021)	(0.0028)	(0.0040)	(0.0049)	(0.0035)	(0.0026)
	γ_{76}	0.0047	0.0037	0.0025	0.0077	0.0081	-0.0012
		(0.0030)	(0.0038)	(0.0045)	(0.0064)	(0.0049)	(0.0038)
	γ_{77}	0.03671	0.0090	0.0342	-0.0059	0.0208	0.0692
		(0.0112)	(0.0122)	(0.0209)	(0.0189)	(0.0174)	(0.0201)
Expenditure-squared	λ_1	-0.0214	-0.0294	0.0138	-0.0182	-0.0267	-0.0024
		(0.0042)	(0.0052)	(0.0075)	(0.0083)	(0.0081)	(0.0088)
	λ_2	-0.0046	0.0109	-0.0402	0.0095	0.0006	-0.0308
		(0.0039)	(0.0047)	(0.0075)	(0.0072)	(0.0078)	(0.0090)
	λ_3	-0.0078	0.0089	0.0043	0.0029	0.0094	0.0170
		(0.0029)	(0.0035)	(0.0056)	(0.0054)	(0.0059)	(0.0062)
	λ_4	-0.0061	-0.0049	-0.0105	-0.0064	-0.0006	-0.0167
		(0.0023)	(0.0027)	(0.0049)	(0.0036)	(0.0052)	(0.0054)
	λ_5	-0.0002	0.0019	-0.0052	-0.0006	-0.0027	0.0028
		(0.0013)	(0.0016)	(0.0026)	(0.0025)	(0.0019)	(0.0037)
	λ_6	-0.0032	-0.0040	-0.0039	-0.0064	-0.0043	0.0024
		(0.0012)	(0.0014)	(0.0017)	(0.0025)	(0.0019)	(0.0018)
	λ_7	0.0276	0.0167	0.0418	0.0192	0.0243	0.0276
		(0.0036)	(0.0038)	(0.0085)	(0.0061)	(0.0064)	(0.0096)
Demographic variables significant?	z's	yes	yes	yes	yes	yes	yes

[1] Standard errors in parentheses

[2] The commodities represented by the different equation numbers are: 1 = grains; 2 = meat and fish; 3 = fruits and vegetables; 4 = dairy; 5 = oils, butter, and other fats; 6 = sugar; 7 = other foods

REFERENCES

Abdulai, A. and D. Aubert "A cross-section analysis of household demand for food and nutrients in Tanzania." *Agricultural Economics* 31(2004):67-79

Agbola, F.W. "Estimation of Food Demand Patterns in South Africa Based on a Survey of Households." *Journal of Agricultural and Applied Economics* 35(3)(December 2003):662-670

Banks, J.; Blundell, R., and A. Lewbel. "Quadratic Engel Curves and Consumer Demand." *The Review of Economics and Statistics* 79(4)(November 1997):527-539.

Blundell, R.W., P. Pashardes, and G. Weber. "What Do We Learn About Consumer Demand Patterns from Micro Data." *American Economic Review* 83(3) (June 1993):570-597.

Blundell, R.W. and J.M. Robin. "Estimation in Large and Disaggregated Demand Systems: An Estimator for Conditionally Linear Systems." *Journal of Applied Econometrics* 14(1999):209-232.

Brownig, M.J., and C. Meghir. "The Effects of Male and Female Labour Supply on Commodity Demands." *Econometrica* 59 (1991):925-951.

Carter, M.R., and J. May. "Poverty, Livelihood and Class in Rural South Africa." *World Development* 27(1)(January 1999):1-20.

Carter, M.R., and J. May. "One Kind of Freedom: Poverty Dynamics in Post-Apartheid South Africa." *World Development* 29(12)(December 2001):1987-2006.

Deaton, A., and J. Muellbauer. "An Almost Ideal Demand System." *American Economic Review* 70(3)(June 1980b):312-336.

Dhar, T., J.P. Chavas, and B.W. Gould. "An Empirical Assessment of Endogeneity Issues in Demand Analysis for Differentiated Products." *American Journal of Agricultural Economics* 85(3)(August 2003):605-617.

Elliot, M.B., and J. van Zyl. "The effect of income on the market for carbohydrates in South Africa." *Agrekon* 30(4)(1991):269-71.

Gould, W. and H.J. Villarreal. "An Assessment of the Current Structure of Food Demand in Urban China." *Agricultural Economics* 34(2006):1-16.

Hancock, P.J., W. L Nieuwoudt, and M.C. Lyne. "Demand analysis of meats in South Africa." *Agrekon* 23(3)(1984):26–29.

Heien, D., and C.R. Wessells. “Demand System Estimation with Microdata: A Censored Regression Approach.” *Journal of Business and Statistics* 8(3)(July 1990):365-371.

Hausman, J.A. “Specification Tests in Econometrics.” *Econometrica* 46(1978):1251-1271

Human Science Research Council (January 2004). *Food Security in South Africa: Key Policies Issues for the Medium Term* (Online). Available from: www.sarpn.org.za/documents/d0000685/Food_security_SA_January2004.pdf

Integrated Food Security Strategy (IFSS). National Department of Agriculture Policy Document. Pretoria, South Africa, 2002.

Jorgenson, D.W., L.J. Lau, and T.M. Stoker. “The Transcendental Logarithmic Model of Aggregate Consumer Behavior.” In R. Basman and G. Rhodes eds., *Advances in Econometrics*, vol.1, Greenwich, Ct: JAI Press, 1982.

LaFrance, J.T. “When is Expenditure ‘Exogenous’ in Separable Demand Models?” *Western Journal of Agricultural Economics* 16(1991):49-62

Leibbrandt, M., and I. Woolard. “A comparison of poverty in South Africa’s nine provinces.” *Development Southern Africa* 16(1)(March 1999):37-35.

Lesser, C.E.V. “Forms of Engel Functions.” *Econometrica* 31(4)(October 1963): 694-703

Lewbel, A. “The Rank of Demand Systems: Theory and Nonparametric Estimation.” *Econometrica* 59(1991):711-730.

Loubser, M. (1990). “Income Elasticities of the demand for consumer goods and services.” Bureau of Market Research No. 175, University of South Africa, Pretoria.

Meenakshi, J.V., and R. Ray. “Regional Differences in India’s Food and Expenditure Pattern: A Complete Demand Systems Approach.” *Journal of International Development* 11(Jan/Feb 1999): 47-74.

Meyerhoefer, C.D., C.K. Ranney, and D.E. Sahn. “Consistent Estimation of Censored Demand Systems using Panel Data.” *American Journal of Agricultural Economics* 87(3)(August 2005):660-672.

Moschini, G. “Units of Measurement and the Stone Index in Demand System Estimation.” *American Journal of Agricultural Economics* 77(February 1995):63-68.

Moro, D. and P. Sckokai. “Heterogenous Preferences in Household Food Consumption in Italy. *European Review of Agricultural Economics* 27(3)(2000):305-323.

Muellbauer, J. “Community Preferences and the Representative Consumer.” *Econometrica* 44(1976): 525-543.

Nieuwoudt, W.L. “The demand for livestock products in South Africa for 2000, 2010 and 2020: Part 1." *Agrekon* 37(2)(1998).

Poi, B. P. “From the Help Desk: Demand System Estimation.” *Stata Journal* 2(4)(2002):403-410.

Pollak, R. A., and T.J. Wales. “Estimation of the Linear Expenditure System.” *Econometrica* 37(4)(October 1969): 611-627.

Pollak, R.A., and T.J. Wales. “Estimation of Complete Demand Systems from Household Budget Data: The Linear and Quadratic Expenditure Systems.” *American Economic Review* 68(3)(June 1978):348-359.

Pollak, R.A. and T.J. Wales. “Demographic Variables in Demand Analysis.” *Econometrica* 49(6)(November 1981):1533-1551.

Poonyth, D., R. Hassan, and J.F. Kirsten. “Random Coefficients Analysis of Changes in Meat Consumption Preferences in South Africa.” *Agrekon* 40 (2001):426-437

Shonkwiler, J.S., and S.T. Yen. “Two-Step Estimation of a Censored System of Equations.” *American Journal of Agricultural Economics* 81(4)(November 1999):972-982.

Taljaard, P.R., Z.G. Alemu, and H.D. van Schalkwyk. “The Demand for Meat in South Africa: An Almost Ideal Estimation.” *Agrekon* 43(4)(December 2004):430-443.

Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press, 2002.

World Bank. GNI Per Capita 2004, Atlas Method and PPP. At: www.worldbank.org

World Development Report (2002), The World Bank.

World Health Organizations (1995). Website: www.who.int