Fractional Multinomial Response Models with an Application to Expenditure Shares

Steven F. Koch
University of Pretoria

27 - 29 October 2010
Indaba Hotel and Conference Centre
Johannesburg
South Africa

Hosted by

DPRU Development Policy Research Unit

TIPS

FRACTIONAL MULTINOMIAL RESPONSE MODELS WITH AN APPLICATION TO EXPENDITURE SHARES

STEVEN F. KOCH[†]

ABSTRACT. The research presented here considers the performance of the Fractional Multinomial Logit (FMNL) model in explaining expenditure shares using data from the 2005/06 South African Income and Expenditure Survey. The results suggest that the FMNL performs favourably, when the dataset is large enough, but that it does not perform as well, when the dataset is limited. Expenditure elasticities were also estimated, and compared to the expenditure shares from a QUAIDS model. The resulting expenditure shares are fairly similar across model specification; however, the FMNL model does incorporate additional curvature, which is easily observed when comparing the QUAIDS elasticities to the FMNL elasticities.

Date: October 2010.
Key words and phrases. Expenditure Shares, Multinomial Fractional Logit.
Prepared for Development Policy Research Unit Conference held in Johannesburg, South Africa, October 27-29. The results presented here are preliminary. Please, do not quote, without the author's permission.
The author would like to thank John Mullahy for sharing his STATA code and for help in using that code. Any remaining errors are the sole responsibility of the author.

1. Introduction

Recent analysis of South African expenditure data has concentrated on changes in various measures of poverty, and, in other cases, considered welfare, more generally. However, only rare occasions has the analysis focused on the deeper empirical realities associated with the estimation of expenditure shares.

For example, when considering the welfare implications of either the Almost Ideal Demand System, or the Quadratic Almost Ideal Demand System, or any number of the theoretical derivatives of these models, the resulting expenditure shares are not forced to fall on or within the unit interval. Although the systems are, in principle, non-linear, the resulting non-linear functions can theoretically lie anywhere on the real number line, and, therefore, the empirical results are not likely to be generally true, although they are expected to be well-behaved at the mean. Since the empirical results are not likely to be correct for a wider range of the data, use of the estimated parameters for welfare considerations will not provide much insight into household welfare away from the mean.

Given the potential problems associated with the empirical results, this research considers an empirical model requiring, as one of its restrictions, all shares to lie within the unit interval. Further, the shares can be estimated as part of a system of equations, thus, there is potential in the modeling framework to allow for non-independence between shares. Most importantly, estimation within a system forces the marginal effects to cancel, as economic theory requires.

The empirical model considered is an extension of the Fractional Logit Model, referred to here as the Fractional Multinomial Logit (FMNL) Model. With the Multinomial Logit Model, the log-likelihood function requires each category to take on a value of either zero or one. With expenditure share data, however, each category can take on a value between zero and one. Therefore, the Multinomial Logit Likelihood Function must be extended to allow for continuous, rather than discrete, data. Allowing for continuous categories results in a relatively simple revision of the Multinomial Logit Likelihood Function, and can be estimated fairly quickly.

Although the parameters in the model are somewhat interesting, since they can contribute to our understanding of welfare, we also estimate average partial effects (marginal effects) and bootstrap confidence intervals associated with the partial effects that are robust to potential asymmetries in the data. Finally, conditional moment tests are examined to determine whether the underlying distribution assumption is consistent with the estimated model.

The empirical results, to date, suggest that the empirical model performs very well. We find that income and household composition have important effects on the budget allocation towards various commodity groups. We also find that the underlying distribution assumptions cannot generally be rejected.

The remainder of the paper is organized as follows. Section 2 desceribes the relevant economic theory and the empirical methodology that will be applied in this research. The data used for the analysis is described in Section 3. Section 4 contains the results of the empirical analysis, while Section 5 concludes and provides some thoughts regarding future work.

2. Methodology

2.1. **Economic Theory.** Consider household $i \in \{1,...,N\}$ with members $h \in \{1,...,H\}$, choosing consumption bundles to maximize utility, which is some function of each household member's consumption vector, $\mathbf{q}^h$, the size of the household, H, and each household member's preferences, Ω^h. We define that utility to be $U = U(\{\mathbf{q}^h\}_{h=1}^H; \{\Omega^h\}_{h=1}^H)$, which is maximized subject to a feasible consumption set, given by $\sum_{h=1}^H \sum_{j=1}^J p_j \cdot q_j^h \leq \sum_{h=1}^H y^h$, where $j = \{1,...,J\}$ denotes the type of good, p_j is the price of good j, and y^h is household member h's income.[1] The optimal solution, indirect utility V, which is one measure of welfare, is therefore a function of the exogenous parameters measuring: income, prices and preferences.

$$V = V(\{p_j\}_{j=1}^J, \{y^h\}_{h=1}^H; \{\Omega^h\}_{h=1}^H) \tag{1}$$

[1]The basic structure does not presume a unitary household model; however, the data available to us is measured at the household level, and, therefore, a unitary model is assumed in the analysis.

Although V represents a measure of welfare, it is not possible to estimate an empirical representation of (1), since the dependent variable is not observable. Instead, empirical welfare analysis is based upon the derivative of V with respect to ln prices. As shown by Chavas & Segerson (1987), the ratios of these derivatives yield household expenditure shares, w_{ij}.[2]

$$\frac{\partial V_i / \partial \ln p_j}{\sum_{k=1}^{J} \partial V_i / \partial \ln p_k} = w_{ij}\left(\mathbf{p}, \mathbf{y}; \mathbf{\Omega}\right) \tag{2}$$

In addition to the preceding result, the budget constraint imposes two further properties on w_{ij}. First, $w_{ij} \in [0,1]$ $\forall i \in \{1,...,N\}$ and $\forall j \in \{1,...,J\}$. Second, $\sum_{j=1}^{J} w_{ij} = 1$ $\forall i \in \{1,...,N\}$.[3]

A wide literature exists discussing potential functional forms for V, and various forms of its dual. Deaton & Muellbauer's (1980) Almost Ideal Demand (AID) System, a rank two demand system has been applied in numerous settings. One advantave of the AID system is that it incorporates a formal welfare structure, yet results in Working-Leser expenditure shares, which are linear in the natural log of expenditure.

$$w_{ij} = \alpha_j + \sum_{k=1}^{J} \sum_{\ell=1}^{J} \gamma_{k\ell} \ln p_k \ln p_\ell + \beta_j \left(\ln y_i - \ln P\right) \tag{3}$$

Banks, Blundell & Lewbel's (1997) Quadratic Almost Ideal Demand (QUAID) System, another often applied model, is a rank three demand system. As the name suggests, the system includes a quadratic in $\ln y$.

$$w_{ij} = \alpha_j + \sum_{k=1}^{J} \sum_{\ell=1}^{J} \gamma_{k\ell} \ln p_k \ln p_\ell + \beta_j \left(\ln y_i - \ln P\right) + \frac{\lambda_j}{b(\mathbf{p})}\left[\ln y_i - \ln P\right]^2 \tag{4}$$

In equations (3) and (4), $\ln P$ is a translog price index, while $b(\mathbf{p})$ is a Cobb-Douglas price aggregator.[4] Some of the variants of the aforementioned models include a

[2]The vector of prices, $\mathbf{p} = \{p_j\}_{j=1}^{J}$, the vector of household incomes, $\mathbf{y} = \{y^h\}_{h=1}^{H}$ and the vector of preferences $\mathbf{\Omega} = \{\Omega^h\}_{h=1}^{H}$ are used to simplify the notation.

[3]A number of other properties, such as Slutsky Symmetry, Cournot Aggregation and Engel Aggregation can also be derived. However, the lack of price data keeps us from empirically considering these properties.

[4]Banks et al. (1997) assume $\ln P = \alpha_0 + \sum_i \alpha_i \ln p_i + 1/2 \sum_j \sum_i \gamma_{ij} \ln p_i \ln p_j$, while $b(\mathbf{p}) = \prod_i p_i^{\beta_i}$.

globally flexible AID system, such as in Chalfant (1987), the modified AID system, as estimated by Fry, Fry & McLaren (1996), and an AID system nested within a Box-Cox structure; see Matsuda (2006).

Empirically, (3) and (4) have been implemented by numerous researchers in even more numerous settings. In the case of South Africa, for example, research using either of these specifications has primarily focused on food demand, as in the analysis of Agbola, Maitra & McLaren (2003), Taljaard, Alemu & van Schalkwyk (2003), Bopape & Myers (2007) and Dunne & Edkins (2008). However, these models have also been used by Koch (2007) to examine alcohol and tobacco expenditure, by Koch & Bosch (2009) to examine the welfare effects of inflation, and by Koch & Alaba (2010) to examine the potential effects of national health insurance on expenditure patterns.

Below, the QUAIDS model is used as an index function within an empirical specification that forces the shares to fall within the unit interval. The QUAIDS specification is then compared to the index model incorporating the QUAIDS specification. Although the analysis is preliminary, the results from the analysis can be used to infer the relative merits of the QUAIDS model in explaining expenditure behaviour in South African households.

2.2. **Econometric Model.** In the preceding subsection, w_{ij} was shown to lie either within or on the boundary of the unit interval, while $\sum_j w_{ij} = 1$. Given this limit, the population model for the share data is assumed to be $E[w_{ij}|\mathbf{z}_i] = G_j(\mathbf{B}, \mathbf{z}_i)$, where $\mathbf{B} = \{\beta_j\}_{j=1}^{J}$. Keeping w_{ij} within the unit interval can be accomplished by assuming the multinomial logit functional form for G_j, using an index function for β_j and $\mathbf{z}_i$, as in (5).

$$E[w_{ij}|\mathbf{z}_i] = G_j(\mathbf{z}_i\mathbf{B}) = \frac{\exp(\mathbf{z}_i\beta_j)}{\sum_{k=1}^{J}\exp(\mathbf{z}_i\beta_k)} \tag{5}$$

The bivariate version of (5), in which $j = 2$, was originally developed by Papke & Wooldridge (1996), who referred to the model as a fractional logit model. The multinomial version, referred to as a fractional multinomial logit model, has been developed, described and applied by Sivakumar & Bhat (2002), Ye & Pendyala (2005) and Mullahy & Robert (forthcoming) for commodity flows, transportation time and household time allocation, respectively.[5] Although the fractional multinomial logit model does not specifically incorporate the unit interval boundary points, except in the limit, modeling either endpoint can be accomplished by including an additional probability function in the population equation to account for the boundary points.[6] However, Mullahy (2010) finds minimal effects of such extensions. Therefore, the model considered here does not further generalize the population mean function beyond (5).

Following Gourieroux, Monfort & Trognon (1984), as was done by Ye & Pendyala (2005) and Mullahy & Robert (forthcoming), we propose a quasi-maximum likelihood (QML) function to simultaneously and efficiently estimate the population equations, assuming the functional specifications in (5) are correct.[7]

$$\mathcal{R} = \prod_{i=1}^{N}\prod_{j=1}^{J} G_j(\mathbf{z}_i\mathbf{B})^{w_{ij}} \tag{6}$$

The natural log of $\mathcal{R}$ is easily constructed, see (7).

$$\begin{aligned}\ln \mathcal{R} &= \sum_{i=1}^{N}\sum_{j=1}^{J} w_{ij}\cdot \ln G_j(\mathbf{z}_i\mathbf{B}) \\ &= \sum_{i=1}^{N}\sum_{j=1}^{J} w_{ij}\left(\mathbf{z}_i' - \ln\left(\sum_{k=1}^{J}\exp(\mathbf{z}_i\beta_k)\right)\right)\end{aligned} \tag{7}$$

Identification of the fractional multinomial model, as in the multinomial logit model, requires normalizing one set of parameters, $\beta_J = \mathbf{0}$, for instance. In other words,

$$E[w_{iJ}|\mathbf{z}_i] = G_J(\mathbf{z}_i\mathbf{B}) = \frac{1}{1+\sum_{k=1}^{J-1}\exp(\mathbf{z}_i\beta_k)} \tag{8}$$

[5]Papke & Wooldridge (2008) have extended the fractional response model to handle panel data.
[6]A tobit-style extension, where $\text{prob}(w_{ij} = 0|\mathbf{z}_i) = \Phi(\mathbf{z}_i\gamma_j)$ could be appended to the model. Mullahy (2010), instead, makes use of the Dirichlet distribution.
[7]If, as is possible, either the index function or the multinimial specification are incorrect, then the estimates will be biased. Future work will consider nonparametric specifications to examine potential specification bias.

and

$$E[w_{ij}|\mathbf{z}_i] = G_j(\mathbf{z}_i\mathbf{B}) = \frac{\exp(\mathbf{z}_i\beta_j)}{1+\sum_{k=1}^{J-1}\exp(\mathbf{z}_i\beta_k)} \quad \forall j \neq J \tag{9}$$

Under the identifaction assumption applied here, r must be modified slightly to account for (8) and (9).[8]

$$\ln\mathcal{R} = \sum_{i=1}^{N}\left[-w_{i1}\cdot\ln\left(1+\sum_{k=1}^{J-1}\exp(\mathbf{z}_i\beta_k)\right)+\sum_{j=1}^{J-1}w_{ij}\left(\mathbf{z}_i'-\ln\left(1+\sum_{k=1}^{J-1}\exp(\mathbf{z}_i\beta_k)\right)\right)\right] \tag{10}$$

The estimated parameters will solve the following first order conditions.

$$\frac{\partial\ln R}{\partial\beta_j} = \sum_{i=1}^{N}\mathbf{z}_i'\left[w_{ij}-G_j(\mathbf{z}_i\hat{\mathbf{B}})\right] = 0 \tag{11}$$

Inference related to the parameter vectors can be calculated using the standard expectations related to m-estimators.[9] Although it is tempting to make use of the underlying multinomial logit distribution to determine the covariance matrix, Mullahy (2010) shows that assumption would lead to underdispersion. Rearranging the m-estimator result, the variance of the parameter estimates is the following:

$$V(\hat{\mathbf{B}}) = \sum_{i=1}^{N}\sum_{j=1}^{J-1}\frac{\mathbf{z}_i'\mathbf{z}_i\left[w_{ij}-G_j(\mathbf{z}_i\hat{\mathbf{B}}))\right]^2}{\left[G_j(\mathbf{z}_i\hat{\mathbf{B}})(1-G_j(\mathbf{z}_i\hat{\mathbf{B}}))\right]^2} \tag{12}$$

2.3. **Partial Effects and Elasticities.** Although parameter estimates are easily calculated through the maximization of (10) and standard errors can be constructed from (12), neither these estimates nor their statistical significance are the primary interest of the analysis. Instead, partial effects, the effect of a change in one of the variables on the expected conditional mean of the share, are the relevant estimates. Partial effects can be calculated from these shares, depending upon whether the variable of interest is discrete or continuous. Assuming a continuous explanatory variable, denoted $z_{i\ell}$, the partial effect of a change in share w_{ij} due to a change in

---

[8]Standard multinomial logit regression packages rely on the presumption that $w_{ij} \in \{0,1\}$. For that reason, the quasi-maximum likelihood function was programmed in STATA's Mata package. Thanks to John Mullahy for sharing his STATA code.

[9]In the case of maximum likelihood, $V(\hat{\theta}) = \mathcal{I}(\hat{\theta})^{-1}\mathcal{S}(\hat{\theta})\mathcal{S}(\hat{\theta})'\mathcal{I}(\hat{\theta})^{-1}$, where $\mathcal{I}$ is the information matrix and $\mathcal{S}$ is the vector of scores.

$z_{i\ell}$ are based on the derivative of the expected conditional mean.

$$\frac{\partial E[w_{ij}|\mathbf{z}_i]}{\partial z_{i\ell}} = \beta_{j\ell} G_j - G_j \sum_{k=1}^{J-1} G_k \beta_{k\ell} \tag{13}$$

Furthermore, since $\beta_{J\ell} = 0$ and $G_j = 1 - \sum_{k=1}^{J-1} G_k$, it is rather straightforward to show that the marginal effects must all cancel.[10]

$$\sum_{j=1}^{J} \left(\frac{\partial E[w_{ij}|\mathbf{z}_i]}{\partial z_{i\ell}} \right) = \sum_{j=1}^{J} \left(\beta_{j\ell} G_j - G_j \sum_{k=1}^{J-1} G_k \beta_{k\ell} \right) = 0 \tag{14}$$

In other words, the multinomal fractional logit model also ensures that the partial effects sum to zero, such that the effect of a change in any variable results in different substitution patterns between goods.[11]

If, instead, the variable of interest is discrete, the partial effect is calculated as a difference. Assuming $\bar{z}_{i\ell}$ is a dummy variable, the discrete partial effect is the conditional mean estimate with the dummy variable turned on net of the conditional mean estimate with the dummy turned off.

$$\frac{\Delta E[w_{ij}|\mathbf{z}_i]}{\Delta \bar{z}_{i\ell}} = G_j(\mathbf{z}_{i\sim\ell} \beta_{j\sim\ell} + \beta_{j\ell}) - G_j(\mathbf{z}_{i\sim\ell} \beta_{j\sim\ell}) \tag{15}$$

In (15), $\mathbf{z}_{i\sim\ell}$ represents the $\mathbf{z}_i$ vector with $\bar{z}_{i\ell}$ removed, $\beta_{j\sim\ell}$ is the conforming parameter vector, and $\beta_{j\ell}$ is the parameter estimate for the discrete variable $\bar{z}_{i\ell}$ for share w_{ij}. As with the continuous partial effect, the sum of all the partial effects associated with the discrete variable $\bar{z}_{i\ell}$ over all shares J, vanishes, yielding pure substitution and complementation effects.

Given the relationships described in (13) and (15), it is readily possible to calculate the partial effects. However, the partial effects do not provide any traction in understanding the significance of any particular partial effect. In principle, the delta method could be used to calculate the standard errors of the marginal effects estimates. However, in this case, we estimate upper and lower confidence values,

---

[10]Mathematically, since $\sum_j w_{ij} = 1 \ \forall i$, the $\sum_j (\partial w_{ij} / \partial z_k) = 0 \ \forall i$ and $\forall k$.

[11]An increase in income can result in an increased share of expenditure on leisure, but, since shares must sum to unity, any increase in one share requires a decrease in at least one other share, although the pattern could be more complex.

based on the description in Hansen (2010). Intuitively, the standard nonparametric bootstrap calculations are most successful when the underlying distribution is symmetric about the true population parameter. Therefore, a generalization to account for possible asymmetry may, in fact, perform much better than the simple bootstrap. By ordering the partial effects estimates, it is possible to extract specific quantiles, such as the 0.975 and 0.025 quantiles associated with a two-tailed test with a 0.05 confidence level, from which confidence intervals can be determined. If the distribution of the estimates is asymmetric, the confidence intervals will reflect that asymmetry by allowing either the upper or lower limit to be nearer to or farther from the average of the estimate.

In the results presented, the focus is upon the Engel curve relationship - the relationship between the natural log of total expenditure and the budget share devoted to each of the particular commodity groups.[12] Given the fact that the partial effects are not linear, and depend upon the level of expenditure, the elasticities are graphed in Figures 1 and 2. The calcuations and graphics were created using R, R Development Core Team (2009).

2.4. **Specification Considerations.** One of the key assumptions underlying all of the previous results, especially the correctness of the reported elasticities, is the specification of the functional form. Functional form is examined via conditional moments tests; equation by equation pseudo-R^2 values are also calculated.[13]

There are a number of ways to consider the appropriateness of functional form in non-linear models. In the context of this paper, we consider a condidtional tests. Since, the first order condition in (11) represents an expectation over a conditional moment, that expectation should hold everywhere in the data. One way to check on the validity of the conditional moment is to consider other functions of $\mathbf{z}$, to see if they are also conditionally uncorrelated with the estimated residuals. Below, we consider quantiles of the fractional mutlnomial regression model to be those functions. Under this construction, we are asking, intuitively, whether or not

[12]The remainder of the partial effects are available from the author upon request.

[13]Future work will consider additional tests as by Ramalho, Ramalho & Murteira (forthcoming).

the moment condition holds across the multinomial logit distribution within each category.

$$C_{\ell q} = N^{-1} \sum_{i=1}^{N} \mathbf{1}\left[G_\ell(\mathbf{z}_i\hat{\mathbf{B}}) \in Q_{q\ell}\right] \times \left(w_{i\ell} - G_\ell(\mathbf{z}_i\hat{\mathbf{B}})\right) \tag{16}$$

In the preceding equation, $Q_{q\ell}$ represents the q^{th} quantile within category ℓ and $\mathbf{1}$ represents an indicator function for whether or not the predicted share lies in the specified quantile.

The last measure of fit to be reported are pseudo-R^2 values. Although R^2 values do not provide any statistical validation of the models, they do provide a rough picture of the capability of the model to explain the variation in the data. The value we report is based on the ratio of the explained variance to the actual variance, and is reported for each share. Assuming there are m variables included in the model, aside from the intercept, and that $\overline{w}_j$ measures the mean of the budget shares in category j, the measure reported is provided in equation (17).

$$\tilde{R}_j^2 = 1 - \sum_{i=1}^{N} \frac{N-1}{N-m} \times \frac{\left(w_{ij} - G_j(\mathbf{z}_i\hat{\mathbf{B}})\right)^2}{\left(w_{ij} - \overline{w}_j\right)^2} \tag{17}$$

Since the QUAIDS model is also reported in the analysis, for comparison purposes, the standard R^2 measures from linear regressions is also reported, share by share, for those estimated equations.

3. Data

The data are taken from the most recent South African Income and Expenditure Survey (IES) conducted by Statistics South Africa (2008). The survey provides information on income, acquisition and expenditure patterns of a nationally representative sample of South African households, as well as fairly detailed sociodemographic information on each individual in the household.[14] Initially, the data was divided into 36 commodities, to match Stats SA CPI data, which is based on the

[14]A household in South Africa includes only individuals who reside in their respective domiciles for at least four days during the week.

Classification of Individual Consumption According to Purpose (COICOP) categories. These 36 expenditure categories were subsequently aggregated into the 12 commodity expenditure groups analyzed here; however, we did not include durable goods in any of our expenditure categories, given the sporadic nature of the purchase of these goods, and the fact that the survey does not include the user cost of durable goods that were purchased in the past.[15]

The IES 2005/2006 is based on both the recall method, for all non-food expenditures, and the diary method, for all food expenditures. A random sample of households was drawn each month, and each household completed both the recall questionnaire and the expenditure diary. Given that the data was completed at 12 different points in time, the data was normalized to March 2006 using the relvant CPI deflator/inflator and aggregated up to incorporate yearly real expenditure. The data has been previously used by Bhorat & van der Westhuizen (2009) to examine changes in poverty and inequality in South Africa between 1995 and 2005, and by Koch & Bosch (2009), who simulate the costs of inflation on household welfare.

From the commodity expenditures, and for estimation of the nonlinear demand system, yearly household expenditure shares were calculated (commodity expenditure divided by total expenditure). Of the 12 shares in our analysis there are four food categories: (1) grain products, (2) protein and dairy products (including meat fish, nuts and oils), (3) fruit and vegetables, and (4) other foods and beverages (including sugar products and candy, coffee and cool drinks). Our commodity shares also include (5) clothing and shoes, (6) housing (including imputed rent), (7) other housing-related consumption (e.g. fuel, power and electricity, and other housing expenses), (8) human capital expenditures (primarily health and education), (9) communications, (10) entertainment (including recreation, reading material, tobacco and alcohol), (11) transport (including public and private transportation costs), and, finaly, (12) other miscellaneous expenditures and investments

[15]Specifically, household total expenditure per annum was calculated to exclude expenditure on vehicles, furniture, appliances, household equipment and textiles, primarily because of their durable nature, while domestic services and other household services were also ignored, given recorded zeroes exceeding 90 per cent of the sample.

Table 1. Expenditure Share Summary Statistics

share	mean	SD	% 0 share	maximum
Grain	0.0828	0.073	1.20	0.68
Proteins	0.1136	0.081	1.62	0.94
Fruit/Veg	0.0382	0.036	4.49	0.56
Food/Bev	0.0619	0.058	2.29	0.64
Clothing	0.0833	0.068	4.07	0.76
House 1	0.1328	0.125	1.66	0.95
House 2	0.1362	0.095	1.08	0.92
Health/Education	0.0347	0.055	9.85	0.93
Communication	0.0356	0.039	10.39	0.74
Entertainment	0.0776	0.098	7.03	0.95
Transport	0.0883	0.087	4.92	0.84
Misc	0.1151	0.114	3.30	0.95

on individuals in the household (including personal care items). Summary statistics for the expenditure share data are presented in Table 1, including the mean, the standard deviation, the percentage of zero shares observed in the data and the maximum share observed in the data.

As can be calculated from the share data in Table 1, the total food share averages approximately 28% of total expenditure, while the other necessities, clothing and housing, cover an additional 35% of total expenditure, leaving 37% of total expenditure allocated to other non-necessities. Further, although no category covers all household expenditure, some households expend a very high proportion of their income on some of the categories. For example, the maximum proportion of expenditure devoted to housing (including imputed rent) and other goods is 95%, while the maximum for many other items exceeds 90%. There are at least some households that do not purchase any of the products contained in each category. Specifically, communication contains the most zero shares, 10.4%, closely followed by health and education, 9.9% and entertainment, 7.0%.

The individual, primarily measured for the household head, and the other household level data used in the analysis are summarized, via means and standard deviations, presented in Table 2. Individual ages are not available in the data, although each individual is placed within an age category. The age category within which the household head falls, is treated as a dummy variable.[16] The age categories were also used to count the number of household members who could be classified as children (aged less than 16), elderly adults (males over the age of 65 and females over the age of 60), or adults (the remaining age groups).[17] The population group of the household head is also treated as a dummy variable. We also include a dummy variable for whether or not the head of the household is a male, and the level of the household head's completed education.[18] The levels of education included are: none, some, primary education, matric, and matric plus.[19] Finally, $\ln x$ and $\ln x^2$ represent the log of total expenditure (net of durable goods and domestic services) and its square, respectively.

Table 2. Explanatory Variable Summary Statistics

Variable	mean	std. dev.	Variable	mean	std. dev.
$\ln x$	10.014	0.96	$\ln x^2$	101.210	19.78
male	0.557	0.50			
age 1	0.057	0.23	black	0.767	0.42
age 2	0.065	0.25	coloured	0.128	0.33
age 3	0.095	0.29	asian	0.016	0.13
age 4	0.112	0.32	white	0.089	0.28
age 5	0.123	0.33	kids	1.332	1.57
age 6	0.112	0.31	adults	2.387	1.56
age 7	0.101	0.30	elders	0.296	0.56
age 8	0.085	0.28	none	0.187	0.39
age 9	0.072	0.26	some	0.314	0.46
age 10	0.063	0.24	primary	0.275	0.45
age 11	0.046	0.21	matric	0.184	0.39
age 12	0.069	0.25	matricp	0.031	0.17

[16]In this footnote, the age categories should be listed.

[17]Males in South Africa are eligible for the old age pension at the age of 65, while females are eligible at the age of 60; thus, the difference in ages used for the classification of elderly adults.

[18]We removed the 173 households, whose head had received a national training certificate (ntc), since these certificates can be earned by people with varying levels of formal schooling, i.e., sometimes ntc is associated with someone who has completed matric and sometimes it is associated with someone who has not completed matric.

[19]Matric is the South African equivalent of high school completion, such that matric plus covers anyone receiving any sort of college education or postgraduate training.

4. Empirical Results

A number of separate models were estimated.[20] Model 1 included $\ln x$ and $\ln x^2$. Model 2 included Model 1, as well as dummy variables for households with a male household head, a white household head, an Asian household head and a coloured household head. Model 3 included Model 2, as well as the number of kids, adults and elderly people. Model 4 included Model 3 as well as categorical variables for the education of the household head: some, primary, matric and matric plus. The final model, Model 5, included Model 4 as well as categorical dummy variables for the age of the household head.[21]

4.1. **Fractional Multinomial Logit Partial Effects.** As has already been noted, the primary concern in the analysis is the average partial effect - the predicted change in the expenditure share due to a change in one of the explanatory variables - as outlined in equations (13) and (15). Given the importance of the relationship between total expenditure and expenditure shares, the focus of the analysis is on the average partial effects associated with expenditure in each of the models.

The estimates of partial effects are included in Table 3 and 4, although only estimates associated with expenditure are included to keep the results parsimonious.[22] Despite the parsimony, there is still a large amount of information contained in the table. Three patterns, however, can generally be discerned in the output. Firstly, the natural log of total expenditure, and its square, are significant and of the same sign for the same goods in all of the model specifications, when significant. Secondly, the natural log of total expenditure, and its sqaure, are insignificant for the same goods across the various specifications. Furthermore, for the most part, the partial effects suggest curvature that is either concave of convex at the mean of

[20]Complete results from all of these models are available from the author upon request.

[21]Although it is tempting to choose the models based on the value of the QML, via a Likelihood Ratio test, such a test is not appropriate. The underlying distribution of the test statistic does not conform to a χ^2, unless the actual QML is a proper likelihood function; see White (1982). Future work will, however, incorporate LM tests for model choice.

[22]Model 5 contains 24 variables across twelve commodities. Therefore, reporting would require three numbers (an estimate, as well as an upper and lower bound) for each of $24 \times 12 = 288$ parameter estimates. All of the partial effects estimates, as well as the parameter estimates, are avaialble from the author upon request.

the data. Thirdly, the average partial effects estimates between Models 1A-1D are largely similar, i.e., the inclusion of household structure only affects the relationship between total expenditure and the budget share associated with certain goods.

Formally, in Models 1A-1D, presented in Table 3, an increase in the natural log of total expenditure is associated with an increase in both grain and protein expenditure; however, that increase is tempered. At the mean of the natural log of expenditure, the average partial effect of (the natural log of) expenditure on both grain and protein shares is negative.[23] Similarly, in Models 1B-1D, communication shares follow a similar pattern. The opposite pattern, shares that decrease at a decreasing rate, is observed for both housing shares, as well as for health and education, while miscellaneous expenditure share partial effects are reasonaby constant, at least at the mean of the partial effects.

Household structure only affects two goods. To see this consider the first two sets of average partial effects estimates presented under Model 1A and Model 1B. In the case of the household's grain budget, the inclusion of household structure - the total number of children, adults and elderly people, each included separately - cuts the first derivative nearly in half. Although the confidence intervals contain some overlap, in both cases, the overlap occurs very close to the average partial effect estimate of the first derivative.[24]

In considering whether or not the FMNL model and the QUAIDS model result in roughly the same estimates, a further analysis was undertaken. In that analysis, a subset of the data was used. The subset was used to keep the data as homogeneous as possible. The partial effects estimates from this further analysis are presented in Table 4. Under QUAIDS, the linear and quadratic terms for the natural log of expenditure are significant determinants of the budget share allocated to protein, other food and beverages, clothing, both housing measures, entertainment and

[23]For grain, for example, the average partial effect is approximately $0.141 - 2 \times 0.009(10.01) = -0.055$ under Model 1A. Even though the model is nonlinear, the average partial effects can be approximated using the derivative of a quadratic function.

[24]For the grain share, the upper limit of the confidence interval associated with $\ln x$ in Model 1B is 0.152, while the mean estimate of the partial effect of $\ln x$ in Model 1A is 0.141.

miscellaneous expenditures. Within the FMNL structure, however, the linear and quadratic terms are only significant within the grain share, protein share, other food and beverages share, and the second housing share. In other words, the QUAIDS model results in more precise estimates, which is not surprising, since the QUAIDS model requires less of the data. In general, however, the confidence intervals from the FMNL average partial effects include the parameter estimates from the QUAIDS estimates, which are the average partial effects, due to the linear structure of the QUAIDS model. The only exception is for the miscellany budget share.

4.2. **Expenditure Elasticities.** The estimates of expenditure elasticitities, on the other hand, are taken from Model 2 estimated, both in Quadratic Almost Ideal Demand System form, as well as in Fractional Multinomial Logit form - using QUAIDS as an index function - from a subsample of the dataset that is assumed to be more homogenous.[25] The estimated partial effects, already discussed, are available for both the QUAIDS and FMNL models in Table 4. Illustrations of the predicted change in expenditure shares across the distribution of the natural log of expenditure are based on these two models, as well.

Below, we look into the effect of $\ln x$ on the predicted share, in order to present a wider view of the expenditure elasticity. Given that there are so many right hand side variables, it was decided to rerun the analysis on a limited dataset. The limited dataset uses only black households with two adults, a male head, three or fewer kids, no elders, while the head only has "some" education.

Figures 1 and 2 provide illustrations of the expenditure elasticity across the distribution of the natural log of expenditure. Although standard errors are not included in the illustrations, it is clear from the illustrations that the QUAIDS model and the FMNL model with QUAIDS index result in rather similar estimates of expenditure share elasticities for most of the goods. Only the grain share elasticity, the first panel of Figure 1 and the miscellaneous budget share elasticity, illustrated

[25]These households are headed by males with some education, black, have two adults, no elders and three or fewer children.

in the last panel of Figure 2 show obvious differences. The primary difference across all of the share elasticity estimates can be understood in the context of the functional forms. Under QUAIDS, the second derivative is expected to be constant, since QUAIDS is a quadratic function; see (4). However, under FMNL with a QUAIDS index, the multinomial logit distribution follows an additional curvature pattern, such that the second derivative is not necessarily constant. Thus, FMNL expenditure elasticity measures tend to include additional curvature, relative to the QUAIDS estimated elasticities.

For grain, protein, other food and beverages, clothing and shoes, and entertainment, expenditure share elasticities are generally decreasing. However, they are not generally always positive or always negative. In the case of the budget allocated to protein, for example, the elasticity is positive at lower levels of total expenditure, but becomes negative at higher levels of expenditure. For the remaining budget shares - fruits and vegetables, both housing shares, communication, transport and miscellaneous goods - elasticities tend to be increasing over the distribution of total expenditure. However, even though the elasticities are increasing, they tend to be nearly always negative or positive for all of these commodities over the distribution of total expenditure.

4.3. **Specification Considerations.** As already noted, the estimated elasticities are quite similar across the model specifications, while the FMNL specification generally requires more data than the linear QUADS specificiation, due to the greater complexity of the model specification. Although there are similarities in the reported results, it is worthwhile considering the performance of the FMNL specification. In order to do so, a conditional moments test, as described in equation (16), was considered. That test was considered both for Model 1D and for Model 2B. The results from the test are presented in Table 5.

Under Model 1D, 13 of the 48 reported results are significantly different from zero, but the conditional moments test does not reject the functional specification ($\chi^2_{36} = 110.88$, $p < 0.0001$). On the other hand, under Model 2B, 0 of the 60 reported results are significantly different from zero, yet the conditional moments test does not accept the functional specification (χ^2_{36} = 47.74, p = 0.483). Once again, the resultsof the conditional moments tests, just like the precision of the average partial effects, point to the need for additional data to improve the precision of the test for Model 2B.

Finally, a comparison of the fitted model R^2 calculations are presented. These are available in Table 6. The results point to rather large differences. The fractional multinomial logit model explains a larger portion of the variation of grain shares, fruit and vegetable shars and miscellaneous shares, while QUAIDS explains a larger proportion of the share variation for the remaining shares, and generally by a rather large margin. Although these calculations do not provide a comparison of the two models, they are indicative, once again, of the need for additional data under the FMNL specification.

5. Discussion and Conclusion

In this paper, we have examined the performance of the fractional multinomial logit model applied to expenditure share data. Although the FMNL does not exactly conform to the behaviour of a utility maximizing household, share data is expected to fall on or within the unit interval. Unfortunately, the usual expenditure share models applied in the literature do not necessarily require estimated shares to fall within the expected interval. Since the FMNL model does include this expectation as a fundamental feature, it may perform better than the models that have been applied in the literature.

The model that was estimated made use of the QUAIDS specification as the index representation within the FMNL structure. The results point to estimates of expenditure elastiticities that are generally very similar to those estimated under the QUAIDS specification, alone. However, additional curvature assumptions are included within the FMNL functional specification, such that the estimated elasticities within the FMNL specification are not linear, as they are in the QUAIDS specification.

The FMNL specification was further examined via conditional moments tests, which, in this case, required quantiles of the distibution to be uncorrelated with the estimated residual. Under the full model specification, in which all households were included, the conditional moments test did not reject the specification. However, when the dataset was limited to a more homogeneous dataset, the conditional moments test did not accept the functional specification, despite the empirical result that no individual conditional moment was estimated to be significantly different from zero. In other words, the conditional moments test for the limited dataset lacked power, due to the reduced number of observations.

The results imply the need to further investigate the performance of the model, possibly through the analysis of fractional regression model tests that have been described by Ramalho et al. (forthcoming). Future work will endeavour to do so.

References

Agbola, F. W., Maitra, P. & McLaren, K. R. (2003), On the estimation of demand systems with a large number of goods: An application to South African food demand, *in* 'Annual Conference of the Agricultural Economics Association of South Africa', Pretoria, South Africa.

Banks, J., Blundell, R. & Lewbel, A. (1997), 'Quadratic engel curves and consumer demand', *Review of Economics and Statistics* **79**(4), 527–539.

Bhorat, H. & van der Westhuizen, C. (2009), Economic growth, poverty and inequality in South Africa: The first decade of democracy, Technical report, The Presidency of South Africa.

Bopape, L. & Myers, R. (2007), Analysis of household demand for food in South Africa: Model selection, expenditure endogeneity, and the influence of sociodemographic effects, *in* 'African Econometrics Society Conference', Cape Town, South Africa.

Chalfant, J. A. (1987), 'A globally flexible almost ideal demand system', *Journal of Business & Economic Statistics* **5**(2), 233–242.

Chavas, J. P. & Segerson, K. (1987), 'Stochastic specification and estimation of share equation systems', *Journal of Econometrics* **35**(2-3), 337–358.

Deaton, A. S. & Muellbauer, J. (1980), 'An almost ideal demand system', *American Economic Review* **70**(?), 338–348.

Dunne, J. P. & Edkins, B. (2008), 'The demand for food in South Africa', *South African Journal of Economics* **76**(1), 104–117.

Fry, J., Fry, T. & McLaren, K. R. (1996), 'The stochastic specification of demand share equations: Restricting budget shares to the unit simplex', *Journal of Econometrics* **73**(?), 377–385.

Gourieroux, C., Monfort, A. & Trognon, A. (1984), 'Pseudo maximum likelihood methods: Theory', *Econometrica* **52**(3), 681–700.

Hansen, B. E. (2010), 'Econometrics', University of Wisconsin Manuscript.
URL: *http://www.ssc.wisc.edu/ bhansen/econometrics/Econometrics.pdf*

Koch, S. F. (2007), 'South African household expenditure shares: South African data pitfalls', *Journal for Studies in Economics and Econometrics* **31**(1), 1–28.

Koch, S. F. & Alaba, O. A. (2010), 'On health insurance and household decisions: A treatment effects analysis', *Social Science & Medicine* **70**(2), 175–182.

Koch, S. F. & Bosch, A. (2009), 'Inflation and the household: Towards a measurement of the welfare costs of inflation', South African Reserve Bank Working Paper WP/09/02.

Matsuda, T. (2006), 'A Box-Cox consumer demand system nesting the almost ideal demand system', *International Economic Review* **47**(3), 937–949.

Mullahy, J. (2010), Multivariate fractional regression estimation of econometric share models. Presented to the 2nd International Health Econometrics Workshop, Rome, Italy.

Mullahy, J. & Robert, S. A. (forthcoming), 'No time to lose: Time constraints and physical activity in the production of health', *Review of the Economics of the Household* **DOI: 10.1007/s11150-010-9091-4**.

Papke, L. E. & Wooldridge, J. M. (1996), 'Econometric methods for fractional response variables with an application to 401 (K) plan participation rates', *Journal of Applied Econometrics* **11**(6), 619–632.

Papke, L. E. & Wooldridge, J. M. (2008), 'Panel data methods for fractional response variables with an applicaiton to test pass rates', *Journal of Econometrics* **145**(1-2), 121–133.

R Development Core Team (2009), *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
URL: *http://www.R-prject.org*

Ramalho, E. A., Ramalho, J. J. S. & Murteira, J. M. R. (forthcoming), 'Alternative estimating and testing empirical strategies for fractional regression models', *Journal of Economic Surveys* .
URL: *DOI: 10.1111/j.1467-6419.2009.00602.x*

Sivakumar, A. & Bhat, C. R. (2002), 'A fractional split distribution model for statewide commodity flow analysis', *Transportation Research Record* **1790**, 80–88.

Statistics South Africa (2008), Income and expenditure of households 2005/6: An analysis of results, Report 01-00-01, Statistics South Africa, Pretoria.

Taljaard, P. R., Alemu, A. G. & van Schalkwyk, H. D. (2003), A linearized almost ideal demand system (la/aids) of the demand for meat in South Africa, *in* 'Annual Conference of the Agricultural Economics Association of South Africa', Pretoria, South Africa.

White, H. (1982), 'Maximum likelihood estimation of misspecified models', *Econometrica* **50**(?), 1–16.

Ye, X. & Pendyala, R. M. (2005), A model of daily time use allocation using the fractional logit methodology, *in* 'Transportation Research Board 2005 Annual Meeting CD-ROM'.

TABLE 3. Average Parial Effects Estimates for Model 1

Variable	Grain	Proteins	Fruit/Veg	Food/Bev	Clothing	House 1	House 2	Health/Ed	Commun	Entertain	Transport	Misc
Model 1A: $QML = -48655.33$												
$\ln x$	0.141	0.155	0.020	0.012	0.028	-0.211	-0.071	-0.051	0.029	-0.098	0.004	0.042
ci-L	0.049	0.116	-0.004	-0.029	0.007	-0.258	-0.113	-0.082	0.006	-0.127	-0.046	0.005
ci-U	0.231	0.194	0.039	0.051	0.047	-0.174	-0.027	-0.028	0.050	-0.062	0.050	0.096
$\ln x^2$	-0.009	-0.009	-0.002	-0.001	-0.001	0.011	0.003	0.003	-0.001	0.005	0.001	0.000
ci-L	-0.014	-0.011	-0.003	-0.003	-0.002	0.009	0.001	0.002	-0.002	0.003	-0.002	-0.002
ci-U	-0.004	-0.007	0.000	0.001	0.000	0.013	0.005	0.004	0.000	0.006	0.003	0.002
Model 1B: $QML = -48442.50$												
$\ln x$	0.088	0.140	0.003	-0.006	0.016	-0.189	-0.087	-0.057	0.043	-0.046	0.037	0.059
ci-L	0.015	0.105	-0.017	-0.042	-0.007	-0.225	-0.127	-0.088	0.021	-0.072	-0.012	0.021
ci-U	0.152	0.181	0.016	0.030	0.036	-0.147	-0.042	-0.034	0.063	-0.016	0.080	0.105
$\ln x^2$	-0.007	-0.008	-0.001	0.000	-0.001	0.010	0.004	0.003	-0.002	0.003	-0.001	0.000
ci-L	-0.010	-0.010	-0.002	-0.002	-0.002	0.008	0.002	0.002	-0.003	0.001	-0.003	-0.002
ci-U	-0.003	-0.006	0.000	0.002	0.000	0.012	0.006	0.005	-0.001	0.004	0.002	0.001
Model 1C: $QML = -48408.42$												
$\ln x$	0.078	0.126	0.003	-0.005	0.021	-0.187	-0.082	-0.056	0.038	-0.041	0.036	0.070
ci-L	0.001	0.087	-0.016	-0.041	-0.002	-0.231	-0.125	-0.086	0.015	-0.068	-0.010	0.032
ci-U	0.133	0.168	0.017	0.034	0.040	-0.156	-0.033	-0.032	0.057	-0.012	0.080	0.124
$\ln x^2$	-0.006	-0.007	-0.001	0.000	-0.001	0.010	0.004	0.003	-0.002	0.002	-0.001	-0.001
ci-L	-0.009	-0.009	-0.002	-0.002	-0.002	0.008	0.001	0.002	-0.003	0.001	-0.003	-0.004
ci-U	-0.002	-0.005	0.000	0.002	0.000	0.012	0.006	0.005	-0.001	0.004	0.002	0.001
Model 1D: $QML = -48359.33$												
$\ln x$	0.079	0.125	0.005	-0.007	0.018	-0.187	-0.080	-0.053	0.037	-0.043	0.036	0.069
ci-L	0.008	0.084	-0.014	-0.045	-0.006	-0.228	-0.121	-0.080	0.013	-0.072	-0.010	0.030
ci-U	0.147	0.166	0.018	0.030	0.038	-0.159	-0.031	-0.031	0.056	-0.014	0.075	0.118
$\ln x^2$	-0.006	-0.007	-0.001	0.000	-0.001	0.010	0.004	0.003	-0.002	0.003	-0.001	-0.001
ci-L	-0.010	-0.009	-0.002	-0.002	-0.002	0.008	0.001	0.002	-0.003	0.001	-0.003	-0.003
ci-U	-0.002	-0.005	0.000	0.002	0.000	0.012	0.006	0.004	-0.001	0.004	0.001	0.001

Source: Output from fractional multinomial logit model using 2005/6 South African Income and Expenditure Survey; $n = 20708$.
ci-U and ci-L: respectively, upper and lower bound for confidence interval of average effect estimate.
Model 1A includes: $\ln x$, $\ln x^2$ and dummies for male household head population group of household head.
Model 1B includes: Model 1A plus kids, adults and elders.
Model 1C includes: Model 1B plus dummies for household head education.
Model 1D includes: Model 1C plus household head age categories.

TABLE 4. Average Partial Effects for Model 2

	Grain	Protein	Fruit/Veg	Food/Bev	Clothing	House 1	House 2	Health/Ed	Commun	Entertain	Transport	Misc
Model 2A: ML = 10307.30												
$\ln x$	0.111	0.360	-0.045	0.147	0.149	-0.124	-0.322	-0.052	-0.041	0.256	-0.194	-0.244
ci-L	-0.016	0.202	-0.113	0.030	0.013	-0.307	-0.510	-0.124	-0.121	0.025	-0.392	-0.457
ci-U	0.239	0.517	0.022	0.264	0.284	0.060	-0.135	0.019	0.039	0.485	0.006	-0.031
$\ln x^2$	-0.008	-0.019	0.001	-0.008	-0.008	0.005	0.016	0.003	0.002	-0.013	0.011	0.016
ci-L	-0.014	-0.027	-0.002	-0.014	-0.015	-0.004	0.007	-0.001	-0.002	-0.025	0.001	0.005
ci-U	-0.001	-0.011	0.005	-0.002	-0.001	0.015	0.026	0.007	0.007	-0.001	0.021	0.027
Model 2B: $QML = -1626.79$												
$\ln x$	0.247	0.327	-0.014	0.143	0.083	-0.179	-0.433	-0.071	-0.072	0.188	-0.230	0.012
ci-L	0.093	0.134	-0.091	0.008	-0.097	-0.380	-0.871	-0.151	-0.173	-0.045	-0.401	-0.237
ci-U	0.392	0.528	0.057	0.275	0.251	0.020	-0.101	0.009	0.031	0.409	0.003	0.263
$\ln x^2$	-0.015	-0.017	0.000	-0.008	-0.004	0.008	0.022	0.004	0.004	-0.010	0.013	0.003
ci-L	-0.022	-0.028	-0.004	-0.015	-0.013	-0.002	0.005	0.000	-0.001	-0.021	0.001	-0.010
ci-U	-0.007	-0.007	0.004	-0.001	0.005	0.019	0.046	0.008	0.009	0.003	0.022	0.015

Source: Output from fractional multinomial logit model using 2005/6 South African Income and Expenditure Survey.
ci-U and ci-L: respectively, upper and lower bound for confidence interval of average effect estimate.
Models include: $\ln x$, $\ln x^2$ using limited subsample; n = 685.
The subsample: Black households with two adults, three or fewer kids, no elders and head has only some education.
Model 2A uses QAIDS; see equation (4), while Model 2B uses FMNL with QAIDS index.

TABLE 5. Conditional Moment Test

Quantile	Grain	Protein	Fruit/Veg	Food/Bev	Clothing	House1	House2	Health/Ed	Commun	Entertain	Transport	Misc
Model 1D: $\chi^2_{36} = 110.88$, $p < 0.0001$												
Quantile 1	-8.98	-3.25	-1.54	-4.83	-16.22	0.46	2.74	0.79	-5.16	16.33	5.96	7.25
ci-L	-16.63	-13.06	-4.94	-9.82	-25.80	-12.41	-8.13	-7.48	-8.02	7.42	-7.04	-2.80
ci-U	-3.09	8.60	1.00	1.03	-6.18	13.35	14.40	10.70	-1.98	26.27	14.85	20.92
Quantile 2	-8.68	5.19	-3.89	-0.90	11.13	-17.31	0.23	6.03	3.92	-33.00	-0.36	-16.18
ci-L	-16.80	-8.95	-9.17	-10.03	-1.67	-35.53	-12.02	-2.29	-0.86	-45.05	-14.38	-30.61
ci-U	-0.41	16.92	1.64	4.49	21.28	-0.63	14.74	11.99	10.00	-18.46	16.29	0.04
Quantile 3	15.38	-1.59	5.00	0.77	13.27	-2.18	1.25	-1.50	5.08	-13.30	-17.02	-11.30
ci-L	5.56	-15.70	-0.08	-3.97	3.67	-22.91	-19.85	-8.82	0.60	-35.51	-25.51	-29.40
ci-U	24.95	6.76	6.66	15.18	24.31	16.81	11.58	11.47	10.72	0.90	3.77	2.83
Quantile 4	2.28	-0.35	0.43	4.97	-8.19	19.03	-4.22	-5.32	-3.85	29.98	11.43	20.24
ci-L	-4.22	-8.02	-1.51	-2.98	-17.55	0.72	-12.58	-19.12	-9.09	15.64	-5.47	7.46
ci-U	9.69	12.57	6.72	9.89	0.23	40.03	10.59	1.94	-1.31	45.78	22.04	31.20
Model 2B: $\chi^2_{48} = 47.74$, $p = 0.483$												
Quantile 1	-0.12	1.50	0.13	0.41	-0.58	0.06	1.53	0.18	-0.44	-0.09	-0.52	0.07
ci-L	-0.62	-0.99	-0.22	-0.42	-1.74	-2.27	-1.40	-0.61	-1.08	-1.24	-2.91	-0.63
ci_U	0.60	2.15	0.42	1.18	1.11	1.13	2.90	1.00	0.51	1.88	0.43	1.11
Quantile 2	0.22	-1.29	-0.10	-0.42	0.71	-2.04	-0.73	0.27	-0.42	1.82	-0.54	0.53
ci-L	-0.82	-2.41	-0.64	-1.60	-1.64	-3.57	-1.69	-0.46	-1.19	-2.90	-1.96	-1.18
ci_U	1.27	2.83	0.48	1.01	2.42	0.09	3.39	1.14	0.82	3.98	2.15	2.11
Quantile 3	0.12	0.51	0.39	0.74	0.49	1.01	0.29	0.43	0.07	-1.68	0.25	-0.13
ci-L	-1.59	-2.02	-0.27	-1.04	-2.28	-0.95	-1.99	-0.60	-0.96	-4.03	-1.58	-2.17
ci-U	1.07	2.09	1.04	1.78	2.38	2.98	2.35	1.45	1.02	2.26	2.74	1.74
Quantile 4	-0.48	-0.23	-0.16	-0.23	0.56	1.03	0.10	-0.31	1.13	-1.42	0.72	-1.09
ci-L	-1.72	-2.05	-0.81	-1.59	-1.64	-0.85	-2.20	-1.02	-0.76	-3.55	-1.39	-3.26
ci-U	1.16	1.67	0.57	1.18	1.54	3.56	1.91	0.77	1.96	3.90	2.84	1.00
Quantile 5	0.27	-0.49	-0.27	-0.50	-1.18	-0.06	-1.19	-0.57	-0.34	1.37	0.09	0.62
ci-L	-0.82	-1.83	-0.69	-1.40	-1.96	-1.68	-3.94	-1.12	-0.83	-3.10	-1.48	-1.20
ci-U	1.49	1.38	0.20	1.02	1.98	0.95	0.97	0.19	0.72	3.54	1.71	2.85

ci-U and ci-L: respectively, upper and lower bound for confidence interval of estimate.

TABLE 6. Expenditure Share R^2 Measures

share	QUAIDS R^2	FMNL $\tilde{R}^2$
Grain	0.1274	0.1345
Proteins	0.0786	0.0330
Fruit/Veg	0.0337	0.1059
Food/Bev	0.0583	0.0268
Clothing	0.0679	0.0088
House 1	0.0916	0.0234
House 2	0.0938	0.0128
Health/Education	0.0358	0.0111
Communication	0.0399	0.0093
Entertainment	0.1147	0.0100
Transport	0.0995	0.0370
Miscellaneous	0.1063	0.1519

† PROFESSOR AND HEAD, DEPARTMENT OF ECONOMICS, DIRECTOR, HEALTH AND DEVELOPMENT POLICY RESEARCH GROUP, UNIVERSITY OF PRETORIA, PRETORIA, REPUBLIC OF SOUTH AFRICA; (O) 27-12-420-5285, (F) 27-12-362-5207.

E-mail address: steve.koch@up.ac.za

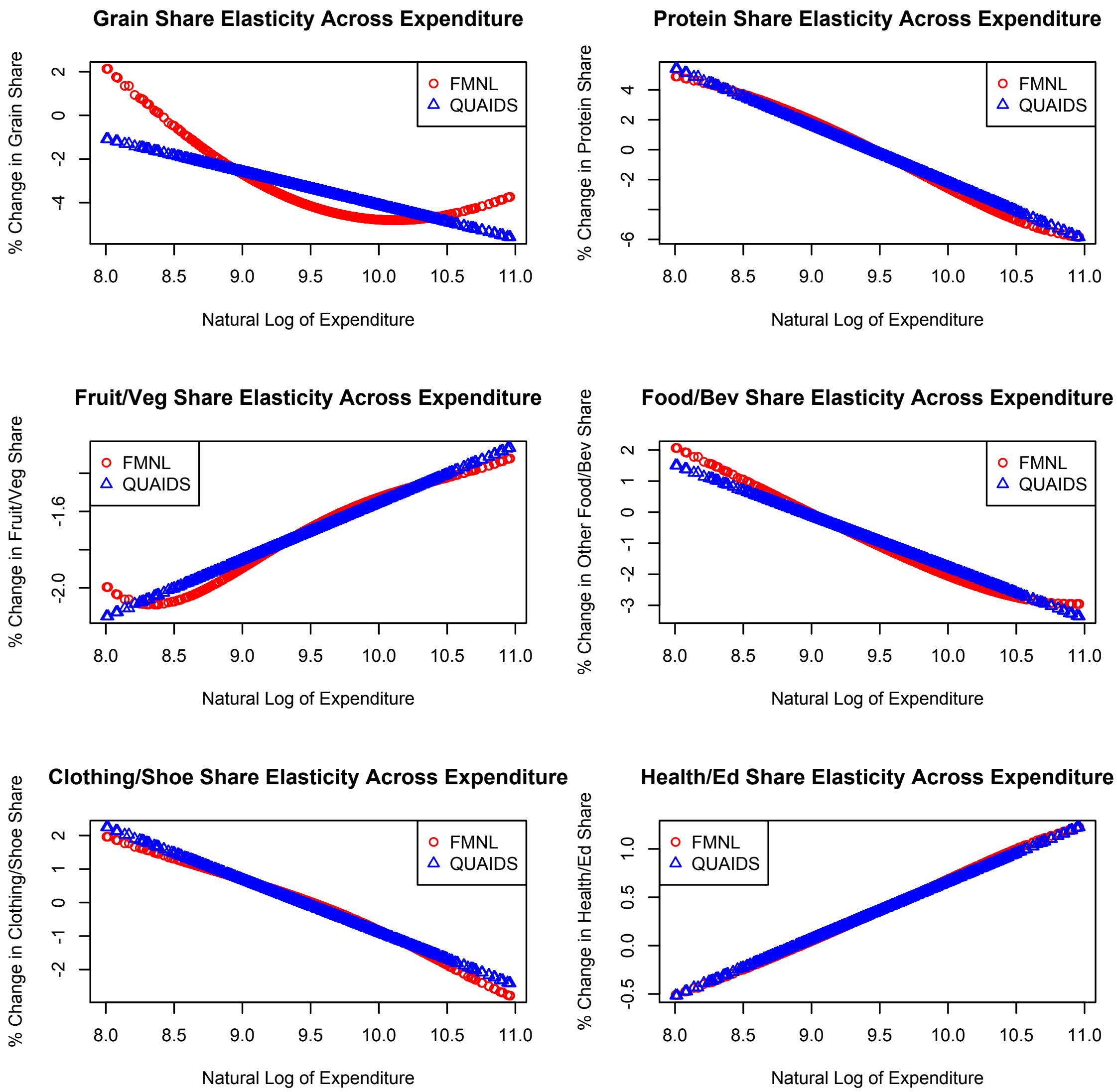

FIGURE 1. Share Elasticities Part I

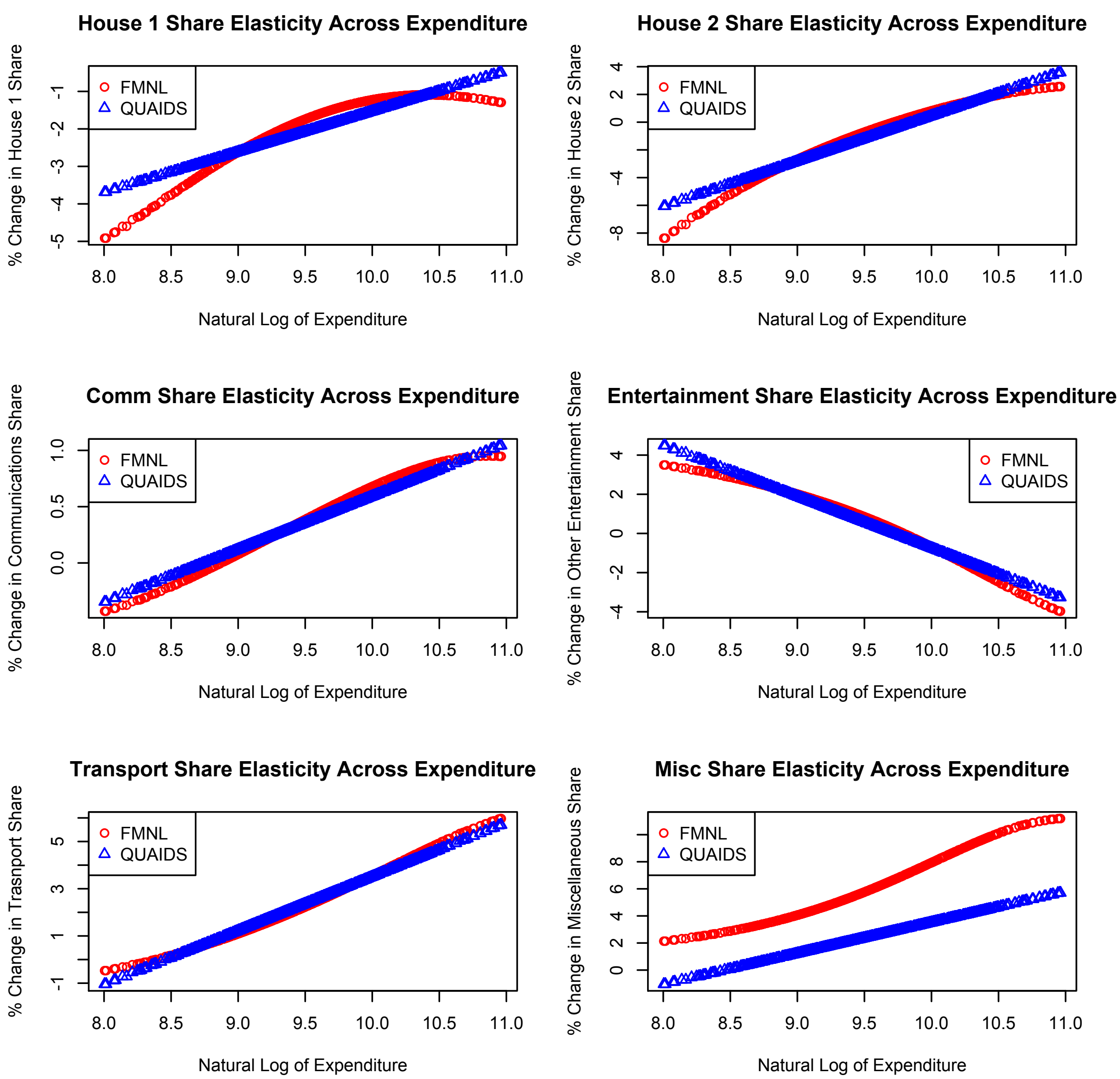

FIGURE 2. Share Elasticities Part II.