

On the Performance of Fractional Multinomial Response Models for Estimating Engel Curves*

Steven F. Koch[†]

June 2013

Abstract

Engel curves are often estimated within a linear, or at least approximately linear system of equations. However, Engel curves are required to lie on or within the unit interval, while summing to unity. These restrictions are not easily accommodated within a linear system. Therefore, we apply the fractional multinomial logit model in our estimation of expenditure shares, because it more readily accommodates these theoretical restrictions. Within our estimation subsample from the South African Income and Expenditure Survey, we find that accounting for these restrictions within the fractional multinomial logit model provides a better fit to the data than does the standard linear system.

Keywords: Expenditure Shares, Multinomial Fractional Logit

*The author would like to thank John Mullahy for sharing his STATA code and for help in using that code. The author would also like to thank seminar participants at the University of Pretoria, the University of Cape Town, Fordham University and the Development Policy Research Unit Conference for their comments. Any remaining errors are the sole responsibility of the author.

[†]Professor and Head, Department of Economics, Director, University of Pretoria, Pretoria, Republic of South Africa; (O) 27-12-420-5285, (F) 27-12-691-2749. steve.koch@up.ac.za

1 Introduction

Originally credited to Engel (1857) and Engel (1895), Engel curves have received much attention in the literature. Although the literature in this area is too vast to discuss in detail - an excellent review is available in Barnett & Serletis (2008) - placing Engel curves, or demand systems, more generally, within the context of utility theory and compositional analysis, Aitchison (1986), is most relevant here. Banks, Blundell & Lewbel's (1997) quadratic ideal demand system, in share form, summarized by equation (20) in Lewbel (2003), and reformulated below, represents the foundation of our empirical analysis.

$$w_{ij} = \alpha_{0j} + \alpha_{1j} \ln x_i + \alpha_{2j} (\ln x_i)^2 + \epsilon_{ij} \quad (1)$$

Since(1) does not include price functions, it is a system of $j \in \{1, \dots, J\}$ Engel curves, where j represents the consumption good being examined and w_{ij} represents the i^{th} household's proportion of total expenditure devoted to consumption good j , and is referred to as an expenditure share. Since each share is a fraction of total expenditure, every share must be non-negative and cannot exceed one, i.e., $w_{ij} \in [0, 1]$. Furthermore, since they are all fractions of the total, $\sum_{k=1}^J w_{kj} = 1$. If price data were available for our analysis, the Engel curves could be extended to include functions of prices, in which case, the system of shares would more generally be referred to as a demand system.¹

Empirically, (1) has been implemented by numerous researchers in even more numerous settings. In the case of South Africa, for example, research using this specification 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, Koch (2007) examines alcohol and tobacco expenditure, Koch & Bosch (2009) examine the welfare effects of inflation, and Koch & Alaba (2010) examine the potential effects of national health insurance on expenditure patterns.

Although QUAIDS has been used extensively, there is one shortcoming, in terms

¹Some of the more common demand systems include Deaton & Muellbauer's (1980) almost ideal demand system (AIDS), as well as Banks et al.'s (1997) quadratic almost ideal demand system (QUAIDS), although there are many variants of these systems. Variants of AIDS and QUAIDS include: globally flexible AIDS, Chalfant (1987), modified AIDS, Fry, Fry & McLaren (1996), and AIDS nested within a Box-Cox structure, Matsuda (2006).

of its empirical application. Specifically, expenditure shares must lie on or within the unit interval, while all expenditure shares must sum to unity. To date, the empirical literature on demand systems has, instead, allowed for the estimates to fall outside the unit interval, and interpreted the results to infer the minimum level of expenditure that would be required to yield positive expenditure shares. Fractional demand systems or the stochastic specification of shares within a fractional system, such as those suggested by Cooper & McLaren (1992), Lewbel (1995), Fry et al. (1996), Cooper & McLaren (1996) and Koch (2007), do require shares to be fractions, and, thus, are capable of restricting shares appropriately. However, shares are rarely restricted to be positive in these models, as is necessary, and, therefore, in application, expenditure shares or Engel curves estimated within fractional expenditure systems often fail to lie on or within the unit interval.

The empirical model that we consider is similar to a fractional demand system, as it is also based on ratios; however, it is set within a probability distribution function that more readily imposes the proper restrictions. In our approach, the QUAIDS model is used as an index function within an empirical specification that forces the shares to be positive, so that they fall within the unit interval, while also forcing the adding-up restrictions to hold. The empirical model considered is the fractional multinomial logit model (FMNL); see Papke & Wooldridge (1996) and Mullahy & Robert (2010). To determine the usefulness of the modelling structure, the original QUAIDS specification is then used as an index function within the FMNL, described below.

The performance comparison is based on simple tests of functional form, that are readily available and rather straightforward to implement. Specifically, QUAIDS budget shares are examined both with respect to Ramsey's (1969) Reset Test, and with respect to Ramalho, Ramalho & Murteira's (2011) extension of the non-nested p-test described by Davidson & MacKinnon (1981). This latter test is based on the relative performance of QUAIDS budget shares with respect to FMNL budget shares. Furthermore, the independent performance of the FMNL model is considered within the context of a conditional moments test, akin to the test applied to logistic regression described in Hosmer & Lemeshow (1989).

The empirical comparison of QUAIDS and FMNL budget shares is undertaken using a subsample of South African households from Johannesburg and Pretoria. These two cities are located in the hub of governmental and business activity and are close to each other, such that the prices in the two cities are extremely similar. The empirical results support the contention that the FMNL modeling structure is preferred to the QUAIDS modeling structure, while the FMNL model is not rejected by the conditional moments test, at least for our subsample of South Africans.

The remainder of the paper is organized as follows. Section 2 outlines the relevant economic theory and the empirical methodology applied in the 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

Although initially proposed by Theil (1969), the fractional multinomial model has recently received renewed attention in economics. For example, research by Sivakumar & Bhat (2002), Ye & Pendyala (2005), Papke & Wooldridge (2008), Mullahy & Robert (2010) and Mullahy (2010), for example, have developed and extended Papke & Wooldridge's (1996) fractional logit model to account for multiple categories. Sivakumar & Bhat (2002), as well as, Ye & Pendyala (2005) develop and explore the FMNL, while both Mullahy & Robert (2010) and Mullahy (2010) have considered some very interesting aspects of the performance of these models. The applications of the fractional multinomial logit model have considered time use in the household, household transport time, statewide commodity flows and portfolio allocations, to date, the model has not been applied to household budget share analysis, although Pu, Lan, Chou & Lan (2008) estimate a bivariate fractional logit model, equation by equation, in their analysis of tobacco substitution effects.

2.1 Econometric Model

Earlier, w_{ij} was shown to lie either within or on the boundary of the unit interval, while $\sum_j w_{ij} = 1$. We assume that the population model for the share data 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 is then accomplished by assuming the multinomial logit functional form for G_j , as in (2), using an index function, as in (1).²

$$E[w_{ij}|\mathbf{z}_i] = G_j(\mathbf{z}_i\mathbf{B}) = \frac{\exp(\mathbf{z}_i\beta_j)}{\sum_{k=1}^J \exp(\mathbf{z}_i\beta_k)} \quad (2)$$

The bivariate version of (2), in which $J = 2$, was 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 (2010) for commodity flows, transportation time and household time allocation, respectively. 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. However, Mullahy (2010) finds minimal effects of such extensions. Therefore, the model considered here does not further generalize the population mean function beyond (2).

Following Gourieroux, Monfort & Trognon (1984), as was done by Ye & Pendyala (2005) and Mullahy & Robert (2010), we propose a quasi-maximum likelihood (QML) function to simultaneously and efficiently estimate the population equations, assuming

²Since this formulation is based on the multinomial logit, it must be noted that there is no assumed correlation between the unobserved determinants of shares. For example, if the transportation share was further disaggregated into car and plane shares, as well as red bus and blue bus shares, it is assumed that whether or not the bus shares are aggregated has no impact on the ratio of food and housing shares. Fortunately, such an assumption is reasonable. However, whether or not the red and blue bus shares are aggregated may have an impact on the relative ratio of car and plane shares. In principle, however, this potential violation of the Independence of Irrelevant Alternatives assumption is not likely to be too problematic for two reasons. The first is due to McFadden (1984), who notes that the restriction is most problematic when there dependent variables that vary by outcome with constant parameter estimates, rather than with constant dependent variables and differing parameters; the latter is the case in our analysis. The second is because aggregation and disaggregation in our analysis automatically changes the shares within the system; therefore, regardless of our choice of aggregation, the shares on the left hand side are appropriately measured. In other words, the red bus and blue bus share would, by definition, aggregate to the the total bus share.

the functional specifications in (2) are correct.³

$$\mathcal{R} = \prod_{i=1}^N \prod_{j=1}^J G_j(\mathbf{z}_i \mathbf{B})^{w_{ij}} \quad (3)$$

The natural log of \mathcal{R} is easily constructed, see (4).

$$\begin{aligned} \ln \mathcal{R} &= \sum_{i=1}^N \sum_{j=1}^J w_{ij} \cdot \ln G_j(\mathbf{z}_i \mathbf{B}) \\ &= \sum_{i=1}^N \sum_{j=1}^J w_{ij} \left(\mathbf{z}'_i - \ln \left(\sum_{k=1}^J \exp(\mathbf{z}_i \beta_k) \right) \right) \end{aligned} \quad (4)$$

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,

$$E[w_{iJ} | \mathbf{z}_i] = G_J(\mathbf{z}_i \mathbf{B}) = \frac{1}{1 + \sum_{k=1}^{J-1} \exp(\mathbf{z}_i \beta_k)} \quad (5)$$

and

$$E[w_{ij} | \mathbf{z}_i] = G_j(\mathbf{z}_i \mathbf{B}) = \frac{\exp(\mathbf{z}_i \beta_j)}{1 + \sum_{k=1}^{J-1} \exp(\mathbf{z}_i \beta_k)} \quad \forall j \neq J \quad (6)$$

Under the identification assumption applied here, $\ln \mathcal{R}$ must be modified slightly to account for (5) and (6).

$$\ln \mathcal{R} = \sum_{i=1}^N \left[-w_{i1} \cdot \ln \left(1 + \sum_{k=1}^{J-1} \exp(\mathbf{z}_i \beta_k) \right) + \sum_{j=1}^{J-1} w_{ij} \left(\mathbf{z}'_i - \ln \left(1 + \sum_{k=1}^{J-1} \exp(\mathbf{z}_i \beta_k) \right) \right) \right] \quad (7)$$

The estimated parameters will solve the following first order conditions.

$$\frac{\partial \ln R}{\partial \beta_j} = \sum_{i=1}^N \mathbf{z}'_i [w_{ij} - G_j(\mathbf{z}_i \hat{\mathbf{B}})] = 0 \quad (8)$$

Inference related to the parameter vectors can be calculated using the standard expectations related to m -estimators.⁴ Although it is tempting to make use of the underlying multinomial logit distribution to determine the covariance matrix, Mullahy (2010)

³If, as is possible, either the index function or the multinomial specification are incorrect, then the estimates will be biased. Below, we consider some specification tests, which show that the model is a reasonable empirical specification.

⁴In 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.

shows that assumption would lead to underdispersion. Rearranging the m -estimator result, the variance of the parameter estimates is the following:

$$V(\hat{\mathbf{B}}) = \sum_{i=1}^N \sum_{j=1}^{J-1} \frac{\mathbf{z}'_i \mathbf{z}_i [w_{ij} - G_j(\mathbf{z}_i \hat{\mathbf{B}})]^2}{[G_j(\mathbf{z}_i \hat{\mathbf{B}})(1 - G_j(\mathbf{z}_i \hat{\mathbf{B}}))]^2} \quad (9)$$

2.2 Partial Effects and Elasticities

Although parameter estimates are easily calculated through the maximization of (7) and standard errors can be constructed from (9), 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 $z_{i\ell}$ are based on the derivative of the expected conditional mean.

$$\frac{\partial E[w_{ij}|\mathbf{z}_i]}{\partial z_{i\ell}} = \beta_{j\ell} G_j - G_j \sum_{k=1}^{J-1} G_k \beta_{k\ell} \quad (10)$$

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.⁵

$$\sum_{j=1}^J \left(\frac{\partial E[w_{ij}|\mathbf{z}_i]}{\partial z_{i\ell}} \right) = \sum_{j=1}^J \left(\beta_{j\ell} G_j - G_j \sum_{k=1}^{J-1} G_k \beta_{k\ell} \right) = 0 \quad (11)$$

In other words, the multinomial 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.

Given the relationships described in (10), 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

⁵Mathematically, since $\sum_j w_{ij} = 1 \forall i$, the $\sum_j (\partial w_{ij} / \partial z_k) = 0 \forall i$ and $\forall k$. An increase in income, for example, 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.

used to calculate the standard errors of the marginal effects estimates. However, in this case, we estimate upper and lower confidence values, 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. Given the fact that the partial effects are not linear, and depend upon the level of expenditure, the elasticities are graphed in Figures ?? and ?. The calculations and graphics were created using R, R Development Core Team (2009).

2.3 Specification Considerations

One of the key assumptions underlying the fractional multinomial logit model and the correctness of the reported elasticities, is the correct specification of the functional form. The examination of functional form is undertaken via a conditional moments test, as well as through the application of a non-nested comparison test of functional specification.

There are a number of ways to consider the appropriateness of functional form in linear and non-linear models. The first specification analysis is based on the Ramsey (1969) Reset Test. In the following analysis, the results suggest that the quadratic format of the QUAIDS specification is not preferred to a model that allows for additional polynomials. The second test to be undertaken is an extension of Davidson & MacKinnon's (1981) non-nested specification test; the extension was proposed by Ramalho et al. (2011). That test is further extended within the context of a system of equa-

tions, through the application of a likelihood ratio test. Finally, the underlying FMNL specification is examined via a conditional moments test.

From Ramalho et al. (2011), consider the two potentially valid functional forms noted in the following equation, where $\mathbf{z}\theta$ represents the QUAIDS functional form, while $G_j(\mathbf{z}\mathbf{B})$ represents the FMNL specification.

$$E[w_j|\mathbf{z}] = \mathbf{z}\theta = G_j(\mathbf{z}\mathbf{B}) \quad (12)$$

Although, for discussion, we assume homoskedasticity, we apply a heteroskedasticity-consistent version of the test, we examine the relative performance of the QUAIDS model, relative to the FMNL. Intuitively, it is a test of whether or not the nonlinear specification given by G_j provides additional information over the initial specification, because it is capable of explaining a significant portion of the QUAIDS residual. The test can be implemented through the following auxilliary regression.

$$(w_j - \hat{w}_j) = \mathbf{z}\delta_j + \gamma_j (\hat{G}_j - \hat{w}_j) + \nu_j \quad (13)$$

The test is referred to as a p-test, because the significance of γ_j in the regression determines whether or not the specification G_j is an improvement over the initial specification. Due to the fact that expenditure shares are part of a system, a test of whether $\gamma_j = 0 \forall j \in J$ is calculated from a comparison of the likelihood values from the restricted and unrestricted seemingly unrelated regression system estimates.

As will be shown below, the FMNL specification provides statistically relevant information over the QUAIDS specification. However, it is not completely clear that the FMNL specification is, itself, valid. In order to check the validity of the FMNL model structure, a conditional moments test is applied. Since, the first order condition in (8) represents an expectation over a conditional moment, that expectation should hold everywhere in the data. One way to check 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 multinomial regression model to be those functions. The test was initially proposed by Hosmer &

Lemeshow (1989) for the analysis of logistic regression. The version applied here is based on Mullahy & Robert’s (2010) implementation. Under this construction, the conditional moments test is, intuitively, a test of whether or not the moment condition holds across the entire multinomial logit distribution within each category.

$$C_{q\ell} = N^{-1} \sum_{i=1}^N \mathbf{1} [G_{\ell}(\mathbf{z}_i \hat{\mathbf{B}}) \in Q_{q\ell}] \times [w_{i\ell} - G_{\ell}(\mathbf{z}_i \hat{\mathbf{B}})] \quad (14)$$

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.

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.⁶ Initially, the data was divided into 36 commodities, to match Statistics South Africa CPI data, which is based on the Classification of Individual Consumption According to Purpose (COICOP) categories. These 36 expenditure categories were subsequently aggregated into the ten 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.⁷

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

⁶A household in South Africa includes only individuals who reside in their respective domiciles for at least four days during the week.

⁷Specifically, 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.

the expenditure diary. Given that the survey was completed at 12 different points in time, the data was normalized to March 2006 using the relevant CPI deflator/inflator and aggregated up to incorporate yearly real expenditure. The data has been previously used by Borat & 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 10 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) capital (including housing expenditures such as fuel and electricity, imputed rent, as well as expenditures on health and education), (7) communications, (8) entertainment (including recreation, reading material, tobacco and alcohol), (9) transport (including public and private transportation costs), and, finally, (10) other miscellaneous expenditures and investments on individuals in the household (including personal care items).

Due to the fact that there is no price data, the analysis is limited to the cities of Pretoria, the administrative capital of South Africa, and Johannesburg, the business capital of South Africa. These cities were chosen due to the fact that they are close together - the centres of these cities are approximately 54km (34mi) apart - and that CPI values over the 12 survey months in these cities were identical. Furthermore, households from these cities were chosen to be as homogeneous. Thus, households were limited to have two or fewer adults, three or fewer children, and were not Asian.⁸ Initially, there were 1410 households in Pretoria and Johannesburg; after our sample selection, 927 remained.

Summary statistics for the analysis data are presented in Table 1, including the mean, the variance, the proportion of zero shares observed in the data, the maximum observed

⁸Only 40 Asian, primarily of Indian descent, households existed in the Pretoria and Johannesburg sample of the data.

value in the data and the proportion of shares lying in the unit interval. Summary statistics for the natural log of expenditure and its square, the two control variables used in the analysis, are also provided in this table.

Table 1: Summary Statistics

Share	Mean	Variance	$(w_j = 0)$	Maximum	$(0 < w_j < 1)$
Grain	0.052	0.002	0.00	0.339	1.00
Protein	0.105	0.006	0.01	0.513	0.99
Fruit/Veg	0.027	0.001	0.04	0.148	0.96
Food/Bev	0.047	0.002	0.02	0.446	0.98
Clothing/Shoes	0.073	0.005	0.06	0.521	0.95
Capital	0.310	0.033	0.00	0.884	1.00
Communications	0.036	0.001	0.05	0.211	0.95
Entertainment	0.106	0.012	0.03	0.862	0.97
Transport	0.122	0.010	0.02	0.771	0.98
Miscellaneous	0.123	0.012	0.00	0.951	1.00
$\ln x$	10.409	1.188		14.473	
$(\ln x)^2$	109.541	544.289		209.457	

Unweighted averages for Pretoria and Johannesburg, South Africa, in the IES 2005/06.

The total food share averages approximately 23% of total expenditure, as can be calculated from Table 1. This share is rather low, and is much lower than for the entire IES 2005/06 data, probably due to the higher income/expenditure for this subsample of South African households; the unweighted survey average exceeds 28%. 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 capital is 88%, while the maximum for miscellaneous items exceeds 95%. There are at least some households that do not purchase any of the products contained in each category. Specifically, clothing and shoes contains the most zero shares, 6%, closely followed by communication, 5%, and fruit and vegetables, 4%. Finally, $\ln x$ and $\ln x^2$ represent the log of real total expenditure (net of durable goods and domestic services) and its square, respectively. Average yearly real net total expenditure in Pretoria and Johannesburg, South Africa, was rather low, amounting to ZAR34 787, approximately US\$4 340. However, the unweighted average for the entire country, according to the 2005/06 Income and Expenditure Survey, was much lower, ZAR22 337.

4 Empirical Results

Initially, the QUAIDS and FMNL models were estimated, the results of which are presented in Table 5. Given the differences in functional form and the difficulties in comparing results across the models, expenditure share elasticities from both the QUAIDS and FMNL models are calculated and presented graphically in Figures ?? and ??. Following the initial estimation and presentation of the estimation results, the empirical validity of the QUAIDS model was examined through the application of both a reset test and a non-nested p-test; these results are available in Table 5. Since, as we see below, the FMNL model is preferred to the QUAIDS structure, the empirical validity of the FMNL structure is examined via conditional moments tests, while the results of that test are available in Table 4.

4.1 QUAIDS and FMNL Estimates

The primary empirical results are presented in Table 5, which contains the QUAIDS estimates, the FMNL parameter estimates, and the FMNL average partial effects estimates. The QUAIDS results, presented in Panel A, show that there is marked curvature, being either concave or convex in the natural log of real net expenditure, in all budget shares except for the capital share. Unfortunately, these parameter estimates do not completely describe the average partial effect of expenditure on the budget shares of interest. The FMNL parameter estimates in Panel B, on the other hand, suggest a relationship that could be different than that from the QUAIDS model, since the sign of many of the parameters are rather different; however, the FMNL parameter estimates are not easily interpreted, due to the non-linear nature of the FMNL structure. Given the difficulty in interpreting the FMNL parameter estimates, average partial effects are also presented in Panel C of the table. Combining the results in Panel B and Panel C, we can see that, although the natural log of net real expenditure, and its square, are significant determinants within the FMNL structure for a number of shares, the average partial effects within the FMNL structure are generally insignificant, with respect to expenditure.

Although the table of empirical results provides basic information regarding the

empirical significance of expenditure in explaining household budget shares for this subsample of households in Pretoria and Johannesburg, the QUAIDS parameter estimates are not the average partial effects, while average partial effects mask the potential variability of the relationship across the distribution of expenditure.⁹ In order to compare the two different sets of partial effects, and not limit that comparison to averages, it was decided, instead, to present a graphical representation of the results. Figures ?? and ?? contain those illustrations in elasticity form.¹⁰ By construction, the QUAIDS elasticity is expected to be linear in the natural log of expenditure, and this result is present in the illustrations. However, the FMNL elasticity is not necessarily expected to be linear, a result that also obtains in the illustrations. Despite the fact that the two results could be rather different, the figures suggest that the QUAIDS and FMNL elasticities are very similar. Only the protein expenditure share elasticity and the capital expenditure share elasticity are largely different, although all of the FMNL expenditure share elasticities exhibit more curvature than their respective QUAIDS expenditure share elasticities.¹¹

Regarding the interpretation of the expenditure share elasticities, recall that income elasticities that are negative are inferior goods, income elasticities between zero and one are normal goods and income elasticities that exceed one are luxury goods. According to the elasticities presenting in Figures ?? and ??, the elasticities are not constant and, therefore, goods are not always inferior, normal or luxuries. The majority of food products are inferior across the range of log expenditure, although food and other beverages are shown to be normal at low levels of expenditure. Clothing and shoes, much like other foods and bevarages are generally inferior, but are normal at low levels of the income distribution. On the other hand, both capital goods and miscellaneous goods are estimated to be luxuries, while the remaining goods - communication, entertainment and transport - range from luxuries at low levels of expenditure to inferior at high levels of expenditure.

⁹For the QUAIDS model, the partial effects are given by $\alpha_{1j} + 2\alpha_{2j} \ln x$ - see (1) - while for the FMNL, the partial effect is described in (10).

¹⁰Rather than illustrating the partial effects across the distribution of expenditure, the budget share elasticity with respect to expenditure is illustrated, instead. That elasticity is the predicted percentage change (out of 100) in the share, and it is graphed against the natural log of expenditure.

¹¹Although confidence bands are not drawn around the elasticities, in order to keep the figures clean, the elasticities are generally not statistically different from each other.

4.2 Specification Considerations

Given the similarity of the elasticities, it would appear that the the FMNL structure is not likely to provide additional information, relative to the QUAIDS model. In order to formally test that conjecture, we considered both a reset test and a non-nested p-test. The reset test provides an indication of whether or not there is more curvature in the share elasticities than suggested by the quadratic structure, while the non-nested p-test provides an indication of whether or not the curvature implied in the FMNL structure is able to explain some of the noise that is not explained by the QUAIDS structure. The results of those two tests are presented in Table 5.

Panel A of Table 5 contains the results from the reset test. Each share equation is re-estimated with a further control for the squared value for the predicted share, which allows for the possibility that the expenditure share could vary with both the cube and quartic of the natural log of expenditure rather than just the first and second order terms initially included. The model is re-estimated within a seemingly unrelated regression system, and the results are presented equation by equation and for the system, as a whole. The equation by equation results suggest that only the communication share and the transport share are independent of additional curvature than was initially modeled in the regression. Furthermore, within the entire system, the square of the predicted share cannot be rejected as a relevant explanatory variable. Thus, the QUAIDS model is not accepted as an appropriate model for expenditure shares in this subsample of Pretoria and Johannesburg households.

Panel B of Table 5 contains the results of the non-nested p-test. As with the reset test and the initial QUAIDS model, it is estimated within a seemingly unrelated regression system, the results of which are presented both equation by equation and within the system, as a whole. Although the reset test suggested that the empirical fit for all expenditure shares, other than for the communication and transport shares, could be improved with the inclusion of additional curvature, the non-nested p-test suggests that, in addition to the communication and transport shares, the fruit and vegetable budget share, the grain budget share and the entertainmnet budget share are not improved through the implementation of the FMNL structure. Across the system as whole, how-

ever, the FMNL structure provides additional empirical information on the QUAIDS residual, and, therefore, the QUAIDS model is rejected in favour of the FMNL model.

4.3 FMNL Conditional Moments Test

Given that the FMNL structure is preferred to the QUAIDS structure, at least for this subsample of households in Pretoria and Johannesburg, it was further decided to examine the appropriateness of the FMNL model via a conditional moments test. Four quantiles, or quartiles, were considered for the analysis, and, therefore, the test considers whether or not the FMNL model is a reasonable fit within each of the four quartiles. The results from the conditional moments test are presented in Table 4. The results contain the conditional mean of the FMNL residual in each of the quartiles, as well as the upper and lower confidence intervals, which are also based on Hansen's (2010) C2 bootstrapping method, described above in relation to the average partial effects. Of the 40 estimates presented in the table, only one conditional mean is significantly different from zero - the capital residual in Quantile 2 - a result that could easily have been obtained by chance. Overall, the χ^2 test statistic is insignificant, and, therefore, the null hypothesis that the conditional residual means are independent of the distribution cannot be rejected. In other words, the FMNL functional form is an appropriate fit for the data.

4.4 Sensitivity Analysis

The results presented so far suggest that the fractional multinomial logit model is better at predicting expenditure share models than the quadratic ideal demand system. However, it is possible that the results from the analysis are driven by other forms of misspecification, such as missing variables or inappropriately including zero observations. Therefore, additional sensitivity analysis is undertaken to further test the underlying specifications.

4.4.1 Dropping all Zero Shares

Neither the quadratic ideal demand system nor the fractional multinomial model directly account for zero shares.¹² Although one possibility would be to estimate a tobit style analysis within each system, similar to that suggested by Shonkwiler & Yen (1999), application of that concept to FMNL is beyond the scope of this paper. Instead, all zero shares were dropped from the data - any household recording a zero share in any of the commodity groupings was dropped from the analysis - and the two models were estimated and tested again.

The results of this analysis are presented in Table ???. As with the initial analysis, neither the reset test null hypothesis nor the non-nested p-test hypothesis are accepted. However, fewer individual shares are statistically affected by either the squared prediction or the predicted information from the fractional multinomial logit model. Within the reset test, protein, other foods and beverages, clothing and shoes, capital and communication shares were all significantly correlated with their squared predictions; however, there were also a few changes. In particular, both the grain and entertainment shares were no longer correlated with their squared predictions, while the communication share became negatively correlated, although only marginally so. Overall, too many shares were correlated with their squared predictions ($\chi_9^2 = 39.15, p = 0.000$). In terms of the non-nested p-test, the FMNL provided additional information for explaining the protein, other food and beverages, capital, communication and entertainment shares. For two of these shares, communication and entertainment, this was a change, compared to the initial data analysis, and, in each case, the additional information was negatively correlated with the share. Again, the FMNL provided additional information for too many shares to reject the null hypothesis of no additional information in the system ($\chi_9^2 = 56.87, p = 0.000$).

Intuitively, it is not all that surprising that removing zero shares has not affected the initial conclusion. In neither the the QUAIDS nor the FMNL are zero shares explicitly accounted; therefore, the comparison of the two is primarily driven by the ability of the two models to keep shares within the unit interval, which the FMNL does by assumption,

¹²If there was no expenditure whatsoever in the household, neither the fractional multinomial logit model nor the quadratic ideal demand system could be estimated, as all shares would be undefined.

and to control for additional curvature, which the FMNL also does.

4.4.2 Controlling for other Covariates

Although the underlying data was meant to be as homogeneous as possible, the initial analysis did not control for household structure; rather, only $\ln x$ and $(\ln x)^2$ were included in the analysis. Therefore, the sensitivity of the results to household structure was also analyzed. Since the initial dataset included households that had no more than three children, no more than two adults and differed by race, dummy variables for one child, two children, three children, two adults and the population group of the household head were also included.

The results from this second sensitivity analysis are presented in Table ???. Interestingly, although the null from the reset test cannot be rejected, the null from the non-nested p-test cannot be accepted. In terms of the reset test, only the entertainment share is significantly correlated with its squared prediction, while the system test implies that this correlation is due to chance ($\chi_9^2 = 11.92, p = 0.218$). However, the fractional multinomial logit model provides additional information over the QUAIDS specification, for the grain, protein, food and other beverages and capital shares. The information provided by FMNL was positively correlated to the grain share residual, although it had not been previously. Further, the additional information from FMNL is no longer capable of explaining any of the clothing and shoe QUAIDS residual. Finally, the results from the four shares, that were significantly affected by the FMNL information, is enough to reject the null hypothesis that the FMNL is not a better conditional expectation than the QUAIDS model ($\chi_9^2 = 59.14, p = 0.000$).

This final set of results is suggestive of the ability of the fractional multinomial logit model's ability to explain expenditure share behaviour in this subsample of South African data. Importantly, the reset test can only provide information on unexplained curvature related to variables already included in the model, and, according to the results, once household structure is accounted for, no further nonlinearities are deemed statistically relevant. However, the reset test does not account for other sorts of nonlinearities, and the results from the non-nested p-test statistically validates the need to

account for other sorts of nonlinearities when examining expenditure shares.

5 Discussion and Conclusion

In this paper, we have examined the performance of the fractional multinomial logit model applied to expenditure share data with respect to the performance of the quadratic ideal demand system, which has featured considerably in the literature. Due to the fact that economic theory requires expenditure to fall on or within the unit interval, while all budget shares must sum to unity, the fractional multinomial logit model, which forces shares to lie within the unit interval, may perform better than the models that have been applied in the literature.

The performance comparison was based on a number of different statistical tests of specification, including different specifications and different subsets of the data, as well as a less formal graphical illustration of the budget share elasticities. The data for the analysis, taken from the 2005/06 South African Income and Expenditure Survey, was limited to households in Pretoria and Johannesburg with few children and few adults, in an effort to keep the households as homogeneous as possible.

Although neither the parameter estimates nor the partial effects estimates are obviously comparable across the two empirical models, the graphical illustration of budget share elasticities suggests that functional form bias might not be an acute problem with this data. However, the graphical analysis suggested that, at least for this data, both the protein budget share elasticity and the capital budget share elasticity could be significantly biased, if the additional curvature implied by the fractional multinomial logit model were ignored, in favour of the simple quadratic ideal demand system. The underlying specification tests, and further sensitivity tests, suggest that expenditure shares should include measures of curvature beyond that included in the quadratic ideal demand system. Specifically, only the fruit and vegetables budget share and the transport budget share are statistically supported by the specification tests in all of the analyses. However, the system as a whole, does not receive statistical support under any of the specification tests, while the fractional multinomial logit does receive support, both independently and in comparison with the initial quadratic ideal demand system.

One of the biggest drawbacks to this study is the lack of price data within South Africa. It is, unfortunately, still possible that the specification tests are, instead of directly testing for curvature, picking-up missing information related to prices. However, it is not clear why the correlation between prices and expenditure could oscillate from positive to negative, as would be necessary to generate the sort of non-linearities implied by the results. Furthermore, it is possible that the non-linearities implied by the tests are more complex than have been picked-up in the analysis, and, for that reason, future work should consider nonparametric analysis. At this time, however, restricting the nonparametric model to account for the the adding-up restrictions requires additional analysis that is beyond the scope of this paper.

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.
- Aitchison, J. (1986), *The statistical analysis of compositional data*, Chapman and Hall, New York, NY.
- Banks, J., Blundell, R. & Lewbel, A. (1997), 'Quadratic engel curves and consumer demand', *Review of Economics and Statistics* **79**(4), 527–539.
- Barnett, W. A. & Serletis, A. (2008), 'Consumer preferences and demand systems', *Journal of Econometrics* **147**, 210–224.
- 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.
- Cooper, R. J. & McLaren, K. R. (1992), ‘An empirically oriented demand system with improved regularity properties’, *The Canadian Journal of Economics / Revue canadienne d’Economie* **25**(3), 652–668.
- Cooper, R. J. & McLaren, K. R. (1996), ‘A system of demand equations satisfying effectively global regularity conditions’, *The Review of Economics and Statistics* **78**(2), 359–364.
- Davidson, R. & MacKinnon, J. G. (1981), ‘Several tests for model specification in the presence of alternative hypotheses’, *Econometrica* **49**(3), 781–793.
- 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.
- Engel, E. (1857), ‘Die productions-und consumptionsverhaeltnisse des koenigsreichs sachsen’, *Zeitschrift des Statistischen Bureaus des Koniglich Sachsischen Ministeriums des Inneren*, No 8 und 9.
- Engel, E. (1895), ‘Die lebenskosten Belgischer arbeiter-familien fruher und jetzt’, *International Statistical Institute Bulletin* **9**, 1–74.
- 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>

- Hosmer, D. W. & Lemeshow, S. (1989), *Applied Logistic Regression*, John Wiley & Sons, Inc., New York, NY.
- Koch, S. F. (2007), ‘The AID and MAID systems: South African data pitfalls’, *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.
- Lewbel, A. (1995), ‘Utility functions and global regularity of fractional demand systems’, *International Economic Review* **36**(4), 943–961.
- Lewbel, A. (2003), ‘A rational rank four demand system’, *Journal of Applied Econometrics* **18**, 127–135.
- Matsuda, T. (2006), ‘A Box-Cox consumer demand system nesting the almost ideal demand system’, *International Economic Review* **47**(3), 937–949.
- McFadden, D. L. (1984), *Econometric analysis of qualitative response models*, Vol. 2 of *Handbook of Econometrics*, North-Holland, Amsterdam, chapter 24, pp. 1395–1457.
- 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. (2010), ‘No time to lose: Time constraints and physical activity in the production of health’, *Review of the Economics of the Household* **8**, 409–432.
- 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.
- Pu, C.-Y., Lan, V., Chou, Y.-J. & Lan, C.-F. (2008), ‘The crowding-out effects of tobacco and alcohol where expenditure shares are low: Analyzing expenditure data for taiwan’, *Social Science & Medicine* **66**(9), 1979–1989.
- 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-project.org>
- Ramalho, E. A., Ramalho, J. J. S. & Murteira, J. M. R. (2011), ‘Alternative estimating and testing empirical strategies for fractional regression models’, *Journal of Economic Surveys* **25**(1), 19–68.
URL: DOI: 10.1111/j.1467-6419.2009.00602.x
- Ramsey, J. B. (1969), ‘Tests for specification errors in classical linear least-squares regression analysis’, *Journal of the Royal Statistical Society, Series B* **31**(2), 350–371.
- Shonkwiler, J. S. & Yen, S. T. (1999), ‘Two-step estimation of a censored system of equations’, *American Journal of Agricultural Economics* **81**(4), 972–982.
- 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.

Theil, H. (1969), 'A multinomial extension of the linear logit model', *International Economic Review* **10**, 251–259.

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 2: QUAIDS Estimates and FMNL Estimates and Average Partial Effects

Panel A: Quadratic Almost Ideal Demand System											
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	Miscellany ^a	
$\ln x$	-0.128*** (0.014)	-0.072*** (0.027)	-0.051*** (0.008)	0.041** (0.018)	0.074*** (0.027)	-0.014 (0.065)	0.036*** (0.013)	0.075* (0.044)	0.128*** (0.039)	-0.090*** (0.041)	
$(\ln x)^2$	0.005*** (0.001)	0.002 (0.001)	0.002*** (0.000)	-0.002*** (0.001)	-0.004*** (0.001)	0.004 (0.003)	-0.002*** (0.001)	-0.004* (0.002)	-0.007*** (0.002)	0.006*** (0.002)	

Panel B: Fractional Multinomial Logit Parameter Estimates											
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	Miscellany ^b	
$\ln x$	1.437** (0.65)	1.144* (0.70)	0.009 (0.74)	1.632** (0.77)	1.954** (0.77)	0.364 (0.88)	1.114 (0.76)	0.794 (0.80)	1.352* (0.76)	0 N/A	
$(\ln x)^2$	-0.118*** (0.03)	-0.091*** (0.03)	-0.034 (0.04)	-0.102*** (0.04)	-0.119*** (0.04)	-0.020 (0.04)	-0.065* (0.04)	-0.054 (0.04)	-0.082** (0.04)	0 N/A	

Panel C: Fractional Multinomial Logit Average Partial Effects											
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	Miscellany ^c	
$\ln x$	0.029	0.029	-0.023	0.037	0.081	-0.129	0.011	-0.003	0.064	-0.094	
ci-L	-0.001	-0.034	-0.036	-0.007	0.018	-0.346	-0.020	-0.082	-0.036	-0.274	
ci-U	0.063	0.122	0.012	0.082	0.150	0.090	0.039	0.082	0.139	0.027	
$(\ln x)^2$	-0.003	-0.003	0.001	-0.002	-0.005	0.010	0.000	0.000	-0.003	0.006	
ci-L	-0.005	-0.008	-0.001	-0.004	-0.008	0.000	-0.002	-0.004	-0.007	0.000	
ci-U	-0.002	0.000	0.001	0.000	-0.002	0.020	0.001	0.004	0.002	0.015	

Observations	927	927	927	927	927	927	927	927	927	927	927
Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ci-L and ci-U: Lower and Upper 0.05 confidence intervals.											
^a Base category for Seemingly Unrelated Regression - parameter estimates and standard errors are residuals ($\sum_j w_j = 1$).											
^b Base category for FMNL - parameters set to zero, no standard errors are available.											
^c Base category for FMNL - average partial effects are residuals (see (10) and footnote 5).											

Table 3: Tests of Functional Form for QUAIDS Relative to FMNL

Panel A: Quadratic Almost Ideal Demand System Ramsey Reset Test ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
\hat{w}_j^2	-2.755* (1.65)	-15.418*** (5.49)	-6.478 (7.77)	22.313*** (8.37)	11.567** (4.64)	-11.560*** (2.64)	-15.985 (76.92)	-37.601* (23.04)	-2.828 (5.63)	49.96*** [0.000]

Panel B: Modified p-test of QUAIDS vs. FMNL ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
$(\hat{w}_j - \hat{C}_j)$	0.683 (1.70)	1.537** (0.76)	-0.529 (1.80)	1.889** (0.78)	2.212*** (0.63)	4.848*** (0.95)	-0.773 (1.68)	-4.762 (3.99)	1.431 (1.35)	81.04*** [0.000]

Standard errors, from 1000 bootstrap replications, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. χ^2 p-value in square brackets.

^a $\ln x$ and $(\ln x)^2$ estimates are all insignificant within the non-nested p-test.

^b Miscellany is reference category for Seemingly Unrelated Regression - parameters not estimated.

Table 4: Conditional Moment Test

Quantile	Grain	Protein	Fruit/Veg	Food/Bev	Clothing	Capital	Commun	Entertain	Transport	Miscellany
Quantile 1	-0.102	0.011	0.067	0.215	-0.308	2.056	-0.482	0.274	0.687	-0.784
ci-L	-0.32	-0.80	-0.17	-0.40	-0.92	-0.79	-0.83	-1.68	-1.60	-2.50
ci-U	0.11	0.55	0.24	0.65	0.52	5.20	0.44	2.44	1.83	0.91
Quantile 2	-0.411	-0.630	-0.381	-1.145	-0.794	-6.710	1.062	-1.973	-3.269	0.891
ci-L	-0.97	-1.68	-0.76	-2.07	-2.96	-10.57	-0.25	-4.31	-5.16	-1.84
ci-U	0.20	1.33	0.09	0.02	1.21	-2.89	1.94	1.87	1.07	3.54
Quantile 3	0.873	0.848	0.210	0.968	0.235	3.608	0.308	-0.687	1.607	0.935
ci-L	-0.16	-1.36	-0.40	-0.50	-1.71	-2.26	-0.75	-3.51	-2.26	-1.90
ci-U	1.79	2.36	0.81	1.98	2.58	8.37	1.29	3.69	3.27	3.72
Quantile 4	-0.360	-0.230	0.105	-0.037	0.866	1.046	-0.888	2.386	0.976	-1.042
ci-L	-1.14	-1.53	-0.46	-1.08	-0.86	-2.01	-1.68	-2.42	-1.43	-3.07
ci-U	0.45	1.19	0.62	1.70	2.35	4.77	0.12	4.41	4.53	1.27

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

Table 5: Tests of Functional Form Dropping Zero Shares

Panel A: Quadratic Almost Ideal Demand System Ramsey Reset Test ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
\hat{w}_j^2	-1.576 (1.35)	-16.646*** (4.72)	-4.205 (11.06)	37.574** (15.37)	29.322* (15.62)	10.734*** (2.54)	-170.222* (96.53)	-61.265 (54.98)	-5.855 (5.99)	39.15*** [0.000]
Panel B: Modified p-test of QUAIDS vs. FMNL ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
$(\hat{w}_j - \hat{G}_j)$	0.415 (0.75)	1.968** (0.61)	0.377 (1.03)	2.061** (0.83)	2.245** (0.88)	3.736*** (0.77)	-11.092* (5.92)	-21.781** (8.78)	0.010 (1.67)	56.87*** [0.000]

Standard errors, from 1000 bootstrap replications, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. χ^2 p-value in square brackets.

^a $\ln x$ and $(\ln x)^2$ estimates are all insignificant within the non-nested p-test.

^b Miscellany is reference category for Seemingly Unrelated Regression - parameters not estimated.

Table 6: Tests of Functional Form with Additional Covariates

Panel A: Quadratic Almost Ideal Demand System Ramsey Reset Test ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
\hat{w}_j^2	0.255 (2.48)	-2.906 (2.03)	6.212 (12.55)	3.642 (7.32)	3.744 (6.18)	0.364 (1.21)	11.187 (28.84)	13.821** (5.59)	-5.550 (6.73)	11.92 [0.218]
Panel B: Modified p-test of QUAIDS vs. FMNL ^{a,b}										
VARIABLES	Grain	Protein	Fruit/Veg	Food/Bev	Cloth/Shoe	Capital	Commun	Entertain	Transport	χ^2_9
$(\hat{w}_j - \hat{G}_j)$	0.919*** (0.18)	0.756** (0.38)	0.849 (1.08)	1.694** (0.65)	1.048 (0.64)	2.962** (1.09)	0.394 (0.87)	1.269 (0.87)	-1.062 (1.00)	59.14*** [0.000]

Standard errors, from 1000 bootstrap replications, in parentheses. *** p<0.01, ** p<0.05, * p<0.1. χ^2 p-value in square brackets.

^a $\ln x$ and $(\ln x)^2$ estimates are all insignificant within the non-nested p-test.

^b Miscellany is reference category for Seemingly Unrelated Regression - parameters not estimated.