ASARC WP 2012/15

Engel Curves and Equivalence Scales for Bangladesh

Syed Abul Hasan*

November 2, 2012

ASARC Working Paper Series

Abstract

This paper examines the Engel curve for major expenditure categories and estimates equivalence scales for Bangladesh. We compare Engel curves estimated by semi-parametric techniques to those arising from models based on consumer theory. Our analysis supports the argument that quadratic food Engel curves are a feature of developing countries. Knowledge about the correct specification of the Engel curve has important implications for modelling household responses to negative income shocks.

JEL-Classification: D11; O21

Keywords: Engel Curve; Semi-parametric Estimation; Semi-nonparametric Estimation; Partial Linear Model; Equivalence Scale; Base Independence; Shape Invariance

*We thank Robert Breunig, Mathias Sinning and Gaurab Aryal for helpful comments. All correspondence to Syed Hasan, HW Arndt Building 25a, Research School of Economics, College of Business and Economics, Australian National University, ACT 0200, Australia, email-syed.hasan@anu.edu.au.

1 Introduction

The Engel curve describes how the expenditure on a commodity varies with household income. We investigate household behaviour by estimating Engel curves for Bangladesh paying particular attention to specification issues. We also estimate equivalence scales for Bangladesh, taking into account demographic characteristics and using a flexible functional form. Our analysis provides evidence for quadratic Engel curves for major expenditure categories – including food – in Bangladesh.

Although a complete demand model would also include relative price effects, Engel curves provide useful insight into many aspects of consumer behaviour. First, the Engel curve has important implications for the design of tax policies. For example, in the case of a quadratic food Engel curve, a higher tax on food items compared to non-food items implies a higher proportion of the tax burden that is borne by the low income people. Second, Engel curves permit a study of intra-household disparity with regard to the distribution of resources including the analysis of, for example, the discrimination of women or elderly persons. Third, Engel curves are crucial in estimating the impact of demographic changes on demand. In a growing economy, these estimates assist in forecasting the demand for some important items like food and energy. Fourth, Engel curves provide the basis for the estimation of equivalence scales and thereby permit welfare comparisons between households. It is also useful for poverty estimation because the minimum consumption varies with demographic characteristics of the family. Finally, Engel curves are useful for predicting the change in a country's trade and production pattern. As a result, the formulation of government policy heavily relies on the Engel curve (Deaton and Muellbauer, 1980; Blundell, Duncan, and Pendakur, 1998; Banks, Blundell, and Lewbel, 1997).

An incorrect specification of the Engel curve not only limits its usefulness but also generates misleading outcomes. For several reasons, the curvature of the Engel curve is particularly important for Bangladesh, where a significant proportion of people live with a subsistence level of income. First, the government of Bangladesh is expanding the coverage of the general sales tax (GST) on food and other necessities (Ministry of Finance, 2012, pp-35).[1] Assessment of welfare impacts of events depends on the shape of the Engel curve. The welfare impact on individuals/households with low income is underestimated if the shape of the Engel curve is quadratic but assumed linear.[2] Second, official poverty estimates in Bangladesh are not based on equivalence scales. Incorporating equivalence scales in poverty estimates would make those more suitable for welfare comparisons. Importantly, estimation of equivalence scales depends on the shape of the Engel curves. Third, household income is volatile in Bangladesh and the income effect differs between a linear and a quadratic Engel curve.

Empirical studies based on household data of advanced economies find a combination of linear and quadratic Engel curves for different goods (Bierens and Pott-Buter, 1990; Banks et al., 1997; Blundell et al., 1998). While these studies typically find a linear relationship between share of food expenditure and income, empirical evidence for rural Pakistan suggests a quadratic food Engel curve (Bhalotra and Attfield, 1998). Restriction from theory, as provided in Blundell et al. (1998), however, asks for a slightly modified semiparametric specification for Engel curves used in Bhalotra and Attfield (1998).

Against this background, this paper studies Engel curves for major expenditure categories in Bangladesh, using a semi-parametric partial linear model (PLM). The advantage of the semi-parametric model is that it does not rely on assumptions about functional form and may therefore help to infer the 'true' shape of the Engel curve. We also include demographics in the models for Engel curve ensuring its' consistency with consumer theory. Demographics enter linearly while also rescale total expenditure in our budget share model.

We begin by estimating Engel curves with the semi-parametric model. To address potential restrictions from consumer theory on PLM, we estimate models with equivalent household expenditure. Assuming a flexible functional form, the equivalence scales are identified semi-parametrically. Finally, using equivalised expenditure, we estimate parametric models

[1] GST is known as the Value Added Tax (VAT) in Bangladesh.

[2] However, such an impact works through relative price mechanisms, which is not captured by our model. Banks et al. (1997) provides an example for this case.

employing a control function approach to address for the potential endogeneity of household expenditure.

Addressing the restrictions imposed by consumer theory on the functional form of the Engel curve and using the data for Bangladesh, our study revisits Bhalotra and Attfield (1998) regarding the shape of the Engel curve for developing countries. We also study the shape of the Engel curve for other important expenditure categories. In doing so, we estimate equivalence scales for households with different compositions. Our study provides additional evidence to the argument put forward by Bhalotra and Attfield (1998) – quadratic food Engel curves are a feature of development status. Failing to account for the curvature of the Engel curve results in underestimating the expenditure variability of low income people suffering from a negative income shock.

We organise the article as follows. In section 2, we discuss the literature on the semi-parametric analysis of Engel curves, paying particular attention to developing countries. We also discuss the method we use to incorporate demographics into the Engel curve model. A brief discussion of the data is presented in Section 3. Section 4 deals with methodological issues. The main models are explained in section 5. We discuss the results in section 6. Section 7 concludes.

2 Semi-parametric Analysis and the Engel Curve

2.1 Low-Income Households and the Food Engel Curve

Bhalotra and Attfield (1998), using a semi-parametric technique, find the shape of the food Engel curve of Pakistan to have a quadratic shape. Motivated by Working (1943), who also find non-linearities in food expenditure at lower level of total expenditure, the study argues that the food Engel curve is quadratic in developing countries, characterised by a large number of low-income households. In such countries, at low income/expenditure levels, the expenditure share on food may either remain constant or decrease more slowly as household income increases.

Grigg (1994), in a slightly different context, discuss possible reasons for an increasing share of food expenditure in developing countries. First, low-income households, being incapable of meeting their nutritional requirements, usually spend almost all of their additional income on food, as income goes up. Second, urban households depend on purchased food, while rural households depend on relatively cheaper farm food.[3] This could imply that the expenditure share of food increases as income increases. Third, with increasing income, consumers may shift from cheap staples to costly food items like eggs, milk, fish and meat. Against this backdrop, it appears likely that the food Engel curve is quadratic in developing countries.

2.2 Incorporating Demographic Characteristics in the Engel Curve

Blundell et al. (1998) show that demographics imposes strong restrictions if it enters linearly in a non-linear Engel curve model like PLM. Specifically, the linearity of demographics implies that if one good assumes the shape of Working-Leser, then all goods needs to be Working-Leser. The study also argues that in the non-linear budget share Engel curve, allowing for flexible shapes over categories requires household expenditure to be adjusted with equivalence scales.[4] Blundell et al. (1998) propose a method to rescale household expenditure through a base independent (IB) equivalence scale. Examples of similar applications include Gozalo (1997) and Pendakur (1999). These techniques are however, suitable for small categories of household with respect to size and composition.

The argument in Blundell et al. (1998) for scaling household expenditure in the PLM provides an incentive to reinvestigate the argument for quadratic food Engel curve for de-

[3] Food prices in rural areas are typically low because rural food is less processed and involves less marketing costs. In addition, a significant proportion of rural food consumption is self-produced for which the recorded price is typically low.

[4] Equivalence scale, used to equivalise household expenditure (income), gives the proportion of expenditure (income) a household (with a particular demographic composition) needs to achieve the same welfare level as the reference household (Breunig and Cobb-Clark, 2005). An equivalence scale is known as base independent if it does not vary with the utility level at which the expenditure comparisons are made (Pendakur, 1999).

veloping countries. A procedure for estimating the equivalence scale for multiple household types is proposed in Yatchew, Sun, and Deri (2003). This procedure is particularly suitable for developing countries, which usually exhibit large variation in household size and composition. Our study aims to investigate the Engel curve for Bangladesh which has not been studied before. This will add to an understanding to consumer behaviour in developing countries.

3 Data

3.1 The Household Income and Expenditure Survey (HIES)

The study utilizes data from the 2010 round of Bangladesh Household Income and Expenditure Survey (HIES) – a repeated cross-sectional household survey. The HIES is conducted every five years and is designed to provide detailed information on household composition, income, consumption and other regional and socioeconomic variables. The households surveyed in the HIES 2010 is selected using a two-stage stratified random sampling approach. The total number of households in HIES 2010 is 12,240. The HIES 2010 is characterised by a large number of households and an extensive use of information and communication technology (ICT) aimed at reducing measurement error (Bangladesh Bureau of Statistics, 2012).

In order to estimate Engel curves, we divide household expenditure into six major categories - 1) food, 2) clothing, footwear and cosmetics, 3) transport, 4) education, 5) medical and 6) other. Several items belonging to a lumpy non-consumption category, like Hajj (pilgrimage) expenditure, are excluded from our analysis.

3.2 Descriptive Statistics

Use of household income in our models results in dropping 380 observations with missing income. To allow for sufficient observations for each demographic group, we also drop 1,297 households. Excluded households either include more than four adults or more than four kids. Furthermore, we trim 2.5% of the data at the top and bottom of household expenditure, which excludes 528 observations.[5] This is due to the fact that in the non-parametric analysis, estimates perform poorly at both end of the distribution, as the lack of observations increases the variance and makes the confidence band wider (Cameron and Trivedi, 2005, pp-316). Our final dataset thus includes 10,035 observations. Family composition, presented below, shows that each households type have a reasonable number of observations in our final dataset.

<insert Table 1 here>

Descriptive statistics for the expenditure categories are presented in Table 2. Mean expenditure share on food is 61 percent, a high number, even compared to other developing countries, e.g., 47 percent in South Africa, as in Yatchew et al. (2000). In addition, food expenditure share varies substantially over its' own quantiles.

<insert Table 2 here>

Table 3 presents means of the dependent variables at different total expenditure quintiles. It conveniently shows rapid reduction of food share as we move towards higher quintiles. However, for the rest of the categories, the pattern is not very clear.

<insert Table 3 here>

We present means and standard deviations of independent variables in Table 4. Our data shows the high variability of household income compared to the household expenditure. This may indicate the high likelihood for household income to suffer from measurement error.

<insert Table 4 here>

[5] Our equivalence scales are slightly modified if we use all observation. However, our conclusions are insensitive to the trimming of data.

4 Empirical Strategy

Two modelling practices guide us in the selection of explanatory variables. First, the measurement error in household income in surveys, as in our case, drive some researchers to model Engel curves on total household expenditure.[6] We follow the same. Second, following important studies like Banks et al. (1997) and Bhalotra and Attfield (1998), we model the expenditure share on the logarithm of the explanatory variable, as such models provide a better fit. The Kernel density estimate of household income and household expenditure – both in log and level – support using the log of household expenditure (log expenditure onwards).

One major drawback of the parametric specification is that the functional form of the equation needs to be specified in advance. Most of the time, the basis of choosing the function is not prior knowledge but tractability of the variables. An application of the non-parametric technique can assist in avoiding specification error (Bierens and Pott-Buter, 1990). However, the curse of dimensionality combined with the practical size of surveys usually restricts researchers to estimate non-parametric regression models involving only one independent variable.

One of the most commonly followed and practical way of avoiding dimensionality problems, while using the advantage of non-parametric regression, is the semi-parametric model.[7] The advantage of the semi-parametric model in estimating the Engel curve is that it generates consistent estimates for the covariates even when these covariates are correlated with the variable whose functional form is not known (Bhalotra and Attfield, 1998). Additionally, parametric specifications can be confirmed by estimating semi-parametric models (Breunig and McKibbin, 2012). One particular type of semi-parametric model – the partial linear model (PLM) – is employed when there is a strong rationale for certain regressors to enter

[6] As Engel and Kneip (1996) pointed out, in surveys, household income data and therefore household disposable income measures are far from perfect. The measurement error is severe in agrarian economies (Bhalotra and Attfield, 1998).

[7] A semi-parametric model allows some independent variables to enter parametrically into the model, while others enter non-parametrically.

the model linearly (Blundell and Duncan, 1998, pp-63). Assuming all the geographic and socioeconomic variables enter the model linearly (while the demographics also scale total expenditure), we choose a PLM in which only household expenditure enters non-parametrically.[8]

For our non-parametric estimation, we prefer the bivariate non-parametric local linear regression using the Kernel method. The choice of local linear regression is due to its improved performance at the boundary compared to kernel and spline smoothers as well as for its consistency and optimal conversion rate (Yatchew, 2003). In non-parametric and semi-parametric models, the selection of the appropriate bandwidth is more important than the selection of the Kernel in that the results are much sensitive to the choice of bandwidth compared to the choice of the Kernel (Yatchew, 1998, pp-684). A high bandwidth leads to a large bias with a small variance, while a small bandwidth generates a larg variance with a small bias. Both cases lead to a higher residual sum of squares and thus a higher mean squared error (MSE). The optimal bandwidth minimizes the integrated version of the MSE – MISE. In our semi-parametric models, optimal bandwidths are based on the cross-validation (CV) approach. The approach is asymptotically equivalent to minimizing a discrete sample approximation of MISE (Hardle and Marron, 1985). We use the Epanechnikov kernel, which constitutes the optimal kernel (Cameron and Trivedi, 2005, pp-314).[9]

The Engel curve, when modelled on household expenditure, may suffer from endogeneity, first noted in Summers (1957). Specifically, household expenditure can be endogenous if households jointly decide total household expenditure and expenditure on each category. In addition, unobserved preference heterogeneity, included in errors in Engel curve models, may also contribute in making total household expenditure endogenous. This is because preferences are likely to be correlated with total household expenditure and thus make the latter endogenous.[10]

For the parametric case, the problem of endogeneity is solved in Liviatan (1961) through

[8]The demographic, geographical and socioeconomic variables of our model are associated with household composition, region, labour market status and season.

[9]The semi-parametric model estimation technique is described in Appendix-A.1.

[10]E.g., Naik and Moore (1996) find unobserved heterogeneity important in modelling food consumption.

using household income as an instrument for household expenditure. For semi-parametric case, Blundell et al. (1998) suggests running a parametric regression of the endogenous variable(s) on a set of instruments. Residuals of this regression are then used as an additional covariate in the semi-parametric model. This procedure can generate consistent estimates of the covariates, while the significance of the residuals indicate the presence of endogeneity. In case of our PLM, we follow an improved methodology, outlined in Newey, Powell, and Vella (1999), which allows us to generate residuals through the non-parametric regression.

In our modelling of Engel curve, the notion of shape invariance has important implications. Engel curves, when plotted as expenditure shares against log of income/expenditure, are called shape invariant if for different demographic groups shapes are identical, except only shifts horizontally or vertically. Base independent equivalence scales provide shape invariant Engel curves.[11] We incorporate demographics in the model assuming base independence, as recommended in Blundell et al. (1998) and Blundell, Chen, and Kristensen (2007). For that, we employ the following functional form for our equivalence scale as recommended in Yatchew, Sun, and Deri (2003),

$$\Delta = (A + \beta_2 K)^{\beta_1} \tag{1}$$

where, Δ is the equivalence scale, A is the number of adult and K is the number of kids in the household. The parameter β_1 captures the economy of scale in household expenditure while β_2 captures the impact of kids on family expenditure. A value of 1 for β_1 indicates that doubling both the number of adults and kids in the household requires expenditure to be doubled for maintaining the same level of welfare. Similarly, a value of 1 for β_2 indicates that, to maintain the same level of welfare, kids require exactly the same amount of resources like adults. The proposed equivalence scale is monotone to both A and K and allows us to model Engel curve with larger variations in family composition (Yatchew et al., 2003). This

[11]The reverse, however, is not true, as mentioned in Pendakur (1999, pp-5). Recently, Lewbel (2010) shows that although it is possible to have a shape invariant demand which is not IB, any quadratic or higher order demand system must be based on (transformed) IB utility. So we use both synonymously.

particularly makes this equivalence scale suitable for our analysis.[12] Our estimation of the equivalence scale is also based on the entire expenditure categories, which is a requirement for consistent estimates, but rare to find in earlier studies. We estimate the model through a grid search for both the parameters over the range from 0.1 to 1.0 by an increment of 0.01.

Finally, we perform a specification test following Hardle and Mammen (1993) to test whether the semi-parametric model can be approximated by a quadratic model. However, as Blackorby and Donaldson (1993) mention, when the expenditure share is modelled on log of household expenditure, it is not possible to identify the equivalence scale and its elasticity. Therefore, we are imposing three assumptions on our model – non-linearity of Engel curves, base independence of equivalence scales and functional form of the equivalence scales. Failure to satisfy any of the three assumptions may result in an unsatisfactory test result.

The study suffers from an important limitation. Because of the use of a large dataset and a relatively complex estimation technique, we are unable to test either the assumptions of base independence or the specification for the equivalence scale.[13] An extension of our study may be testing these assumptions with the aid of a simpler non-parametric technique.

5 Model and Estimation

To illustrate our model, let us start with two households, A and B, with the former being the reference household. Let p be the vector of prices common to both households and let u represent some arbitrary level of utility. Then

$$E^A(p,u) = \frac{E^B(p,u)}{\Delta^B(p,u)} \tag{2}$$

where, $E^h(p,u)$ is the expenditure function, which gives household $h \in \{A, B\}$ an utility level u at prices p; $\Delta^B(p,u)$ is the equivalence scale, which translates B's expenditure into

[12] Our final dataset includes a total of 20 different types of household.

[13] Studies like Pendakur (1999) use a smaller data set while Yatchew (2003) follows a simpler non-parametric technique (running mean smoother) allowing to test base independence assumption and equivalence scale function.

an expenditure level that would give A the same utility as B (equivalent expenditure).

Now, under the base independence assumption, which assumes that the equivalence scale does not depend on the utility level, $\Delta^B(p,u) = \Delta^B(p)$. Therefore, Equation(2) can be expressed with the dual indirect utility function as

$$E^B(p, V(p,y)) = \Delta^B(p) E^A(p, V(p,y)), \tag{3}$$

which is equivalent to

$$V^B(p,y) = V^A(p,y)/\Delta^B(p). \tag{4}$$

Since the expenditure functions are homogeneous of degree one in prices, the equivalence scale function must be homogeneous of degree zero in prices. Thus Blackorby and Donaldson (1993) express (4) as

$$V^B(p,y) = V^A\left(p, \frac{y}{\Delta^B(p)}\right). \tag{5}$$

Pendakur (1999), using Roy's identity, derive Marshallian demand equations

$$x_i^B(p,y) = x_i^A\left(p, \frac{y}{\Delta^B(p)}\right)\Delta^B(p) + \frac{y}{\Delta^B(p)}\frac{\partial \Delta^B(p)}{\partial p_i}, \tag{6}$$

where x_i^h represents total expenditure on goods i for household h. Multiplying equation(6) with p_i/y gives us the Marshallian expenditure share equations

$$w_i^B(p,y) = w_i^A\left(p, \frac{y}{\Delta^B(p)}\right) + \eta_i^B(p) \tag{7}$$

where w_i^h gives the expenditure share of commodity i for household h and $\eta_i^B(p)$ is the elasticity of the equivalence scale w.r.t the price of good i for household B. If prices are constant, as in our cross-sectional data, the model reduces to

$$w_i^B(y) = w_i^A\left(\frac{y}{\Delta^B}\right). \tag{8}$$

We also include some non-demographic variables in our model like regional and urban/ rural dummies, work status and educational level of both the household head and his/her partner. Now, expressing expenditure share of a category as a function of log expenditure and generalisation for more household categories gives our model as

$$w_i = F(logy - Z\delta) + V\lambda_i + \upsilon \tag{9}$$

where F is an unknown function, $\delta = log\Delta$, Z is a vector of indicator variables for demographic composition, V includes Z and a set of non-demographic variables enters linearly in the model, λ_i is a vector of parameters and the error term $\upsilon \sim NID(0, \sigma^2)$.

Substituting the equivalence scale Δ from (1) gives our final model,

$$w_i = F\left(logy - Z\beta_1 log(A + \beta_2 K)\right) + V\lambda_i + \upsilon. \tag{10}$$

As recommended in Yatchew et al. (2003), for the equivalence scale function we consider only a subset of Z.

6 Results and Discussion

As a diagnostic check for the shape of the Engel curves, we first perform quantile regressions for all the expenditure categories.[14] Our plots of the conditional quantile estimate of the coefficients of log expenditure (Figure 1) shows that though the direction of the impact is the same across quantiles of the dependent variables, there is a systematic pattern in the variation of the coefficients, indicating towards non-linear Engel curves.

<insert Figure 1 here>

We estimate the expenditure shares using the semi-parametric regression technique with control for the endogeneity of total household expenditure. In doing so, we also identify

[14] As mentioned in Koenker and Hallock (2001) and Knight and Ackerly (2002), quantile regressions are useful for diagnostic checks with heteroscedastic error terms and outliers in the dependent variable.

the equivalence scales considering all six expenditure groups. That is, we choose values for β_1 and β_2 that minimises the grand total for the sum of squared residuals (SSR) for all six groups.[15] Our model indicates a substantial scope for the economies of scale (a value of 0.74 for β_1) and a low impact of kids on household expenditure (a value of 0.17 for β_2). Given the high expenditure share of food and given that food is a rival good, the estimates for the economies of scale in family expenditure seems sensible.[16] In addition, lower food requirements of the kids together with minimal non-food expenditure at low income jointly explains lower value for β_2. The estimated equivalence scales are presented in Table 5.

<insert Table 5 here>

One implication of the base independence assumption is that the semi-parametrically estimated Engel curves for households with different family composition would be of similar shape. Graphically, this indeed holds for food, as presented in Figure 2. The shape of the Engel curves for our selected demographic groups also look similar for both for the clothing and other expenditure categories. For some categories, the Engel curves have different shapes, which might be a consequence of low number of observations. However, expenditure categories that appear to be shape invariant, like food, clothing and other expenditures, altogether account for around 90% of total household expenditure.

<insert Figure 2 here>

To form an idea about the shape of the Engel curves, we present the semi-parametric fit in Figure 3.[17] We compare the shape of the semi-parametric model to a quadratic fit. The figure shows that for some categories including food, the quadratic model fits the data relatively well.

[15] In our analysis, the SSR has a global minimum for each of the parameters, given the value of the other.

[16] Notably, we find that the equivalence scale parameters does not vary much even if we consider only one category – food or a subgroup like food and cloth. This focuses the fact that most of the variations in the equivalence scales comes from food expenditure.

[17] We thank Vincenzo Verardi and Nicolas Debarsy of University of Namur, Belgium, for sharing their Stata codes for semi-parametric model estimation, see Verardi and Debarsy (2011).

<insert Figure 3 here>

We check if the semi-parametric fit can be approximated by a parametric adjustment of order two, by conducting the specification test proposed in Hardle and Mammen (1993). When our choice of the level of significance is 10%, we cannot reject the hypothesis that 'the quadratic fit and the non-parametric fit are not different' for food, clothing, medical and other categories. We are also unable to reject the same hypothesis for the transport expenditure share at 5% level of significance. However, we reject the null hypothesis of quadratic fit for education over 1% level of significance. Rather than the non-linearity of the models, a violation of the base independence assumption or a misspecification of the equivalence scale may be responsible for such rejections.

<insert Table 6 here>

Next, we fit a parametric model which is quadratic with equivalised expenditure and linear in other covariates.[18] In the estimated expenditure share equations, presented in Table 7, both the level and the quadratic terms are significant for food, clothing, transport, education and other items at 5% significance level. As is common in the case with Engel curves, the adjusted R^2, though reasonable for food, education and other items, is particularly low for medical expenditures.

<insert Table 7 here>

It can be argued that the non-linearity of the Engel curves may be due to the specification error. In our models, we allow the demographic and non-demographic variables to shift the curve. However, we also restrict the value of the coefficients across heterogeneous groups. This might results in spurious curvature of the Engel curve. Nonetheless, we find that, for all

[18] The problem of heteroscedasticity is common in the estimation of Engel curves. At higher income levels, people are usually more flexible in allocating resources across categories. This results in a higher variability of expenditure compared to the variability of expenditure at low incomes, when most resources are allocated to subsistence. Since we find the same for our case, we use robust standard errors.

six expenditure categories, the shapes of the Engel curves do not differ much for households across different regions, industries or occupations. Tobit models, applied on categories with a large fraction of missing values, produce similar results.

As discussed earlier, there are arguments for endogeneity of household expenditure in our model of Engel curves. However, in a two-stage budgeting system, total household expenditure may be exogenous in our models of Engel curve. In that process, household first decides income and total expenditure and next, in the second stage, given total expenditure, decides about expenditures on each category. This particularly seems valid for people with subsistence income, as our case is. Because, at low income, people usually do not have much scope to decide their income based on their expenditure, which makes household income exogenous in the model. In addition, low-income households are forced to spend all of their (inadequate) income on consumption. Now, given total expenditure, they decide expenditure on each category, which makes household expenditure exogenous in our models.

Now, we check for the potential endogeneity of household expenditure, as OLS estimates are inconsistent when household expenditure is endogenous while IV estimates are inefficient when household expenditure is exogenous. For that we use household income, its square and household's land ownership status as instruments. It provides evidences of endogenous household expenditure – except education, in Engel curves for all categories, we reject the exogeneity of log expenditure at 10% level of significance (Table 8).

Even with several limitations, instrumental variable (IV) seems a natural candidate to obtain an unbiased estimate of the Engel curve. Indeed, empirical studies of the Engel curve like Banks et al. (1997), Blundell et al. (1998), Bhalotra and Attfield (1998) and Blundell et al. (2007) routinely perform IV estimation to deal with the problem of endogeneity.[19] In our model, in addition to log income and its' square, we use another instrument – household's landholding status – expected to be closely related with household income and expenditure. This allows us to check for the quality of the instruments. Looking at the first stage regres-

[19]Nonetheless, in a non-linear system, endogeneity may also arise if we use instruments for a variable suffering from measurement error. This is quite common in household data.

sion, we find our instruments strong. Next, we check if the instruments satisfy the exclusion restriction. A test of over-identifying restrictions cannot reject the null hypothesis that the excluded instruments are valid for all expenditure categories at the 5% level of significance (Table 8). However, low p-values for three expenditure categories in the over-identifying restrictions test suggest that at least one of our instruments might not be valid.

<insert Table 8 here>

Surprisingly, with IV, only the Engel curve for clothing expenditure category appears quadratic. In all other cases, not only the quadratic term but also the level term appears insignificant (Table 9). An investigation into the first stage reveals that the high collinearity between the two endogenous variables – log expenditure and its square – is responsible for inflating the standard error and thus making estimates insignificant.

<insert Table 9 here>

<insert Figure 4 here>

Another way of dealing with the endogeneity issue is the control function (CF) approach. Compared to IV, the CF approach has an advantage of providing more precise estimates when the endogenous variable enters into the model non-linearly (Wooldridge, 2010). In the CF approach, residuals from the first stage regression(s) enter as additional covariate(s) in the original model and control for endogeneity. Additionally, significance of the first stage residuals can indicate the endogeneity of the suspected variable(s). As we have two endogenous variables in our models (log consumption and its' square), including two reduced form residuals produces estimates identical to the IV approach. However, now we can test whether both of the first stage residuals are contributing in controlling for the endogeneity. In case we find that either one is enough to control the endogeneity, estimates will be more efficient in our non-linear setting. Therefore, we follow the CF approach in estimating our models.

In selecting our preferred set of models, for each expenditure category we compare models that include two reduced form residuals against the models that only include the residuals from the reduced form of the level term. Though Akaike information criterion (AIC) marginally supports bigger model for cloth and transport category, Bayesian Information Criterion (BIC) unambiguously supports the parsimonious models for all six expenditure categories. This reveals that the residuals of the reduced form for log expenditure is enough to control for the endogeneity for all the expenditure categories.[20] These parsimonious models provide our final estimates (Table 10). In these estimates, Engel curves for all the expenditure categories except medical expenses are quadratic. It is therefore consistent with the argument in Bhalotra and Attfield (1998) that developing countries may have a quadratic food Engel curve. Estimated Engel curves with the CF approach are presented in Figure 5.

<insert Table 10 here>

<insert Figure 5 here>

One of the arguments for a quadratic Engel curve is that low-income households spend all of their additional income on food to meet their nutritional requirement. If this argument is valid, we would expect similar slopes for different food categories. Another argument is that with increase in household expenditure, consumers shift from cheap staples to costly food items. This implies changes in the composition of food consumption to a change in household expenditures.

To investigate different arguments for quadratic Engel curve, we use our earlier framework – semi-parametric PLM, but now include disaggregated expenditure categories like rice, protein, fruits & vegetables, non-home-made food, otherfood, tobacco, clothing, footwear & cosmetics, transport, education, medical and other expenditure. In our new setting, the value for the economies of scale parameter in the equivalence scale function is 0.57 – lower than the value from our earlier setting, 0.74. However, the impact of kids on household

[20]Banks et al. (1997) finds the same for the quadratic model.

expenditure remains the same at 0.17. In our analysis, as expected, household expenditure has the opposite impact on the expenditure share of some important food categories like rice and protein. This provides limited support for the first argument but supports the second. As household expenditure increases, expenditure on categories like protein, non-home-made food and otherfood dominates over expenditure on rice and fruits & vegetables, resulting in a quadratic Engel curve. A quadratic food Engel curve may therefore successfully aggregate different types of food expenditures.

Income Shock and Food Consumption of Low Income People

Earlier, we argue that quadratic Engel curves yield different policy implications. Here we provide an example for the case of the food Engel curve for Bangladesh. Let us consider the case of a negative income shock observed by all households in the economy. This allows us to study the impact of such an event on food expenditure of a low-income household, say with expenditure equal to the 10th percentile of our equivalised expenditure.[21] The estimated elasticity of the food share from the quadratic model for the reference household (single adult with no kids) at the sample means of other regressors in the model is -0.17. However, a model linear in household expenditure provides an elasticity equal to -0.23.[22]

Now, a high income elasticity of the food share implies that in case of a negative income shock households will substantially increase their expenditure share on food and thus minimally reduce their expenditure on food. Conversely, a lower income elasticity of the food share, as is the case for our quadratic Engel curve, would indicate a substantial reduction in food expenditure. In other words, the impact of an income loss on the food share is overestimated in linear models resulting in an underestimation of the impact on food expenditure. At an already inadequate income level, this can have severe consequences on the food intake of low-income households. Policies based on linear Engel curves therefore underestimate the

[21] Using the HIES data, the official estimates of poverty is 31.5 percent in Bangladesh.

[22] In the linear model we also use equivalised household expenditure even though the equivalence scale is not identified in such models.

negative impact of such events and ultimately provide inadequate policy support to the poor.

7 Conclusion

We attempt to identify the shape of the Engel curve for major expenditure categories in Bangladesh using recent household data. To determine the shape from the data, we use a semi-parametric technique with base independence assumption and also control for endogeneity. Our estimate of the equivalence scales seems reasonable, while the test for the semi-parametric specification reveals that Engel curves for most of the categories have a quadratic shape. We then estimate quadratic model with equivalised household expenditure and check the significance of the quadratic term. We find that the quadratic term is highly significant for food, clothing, transport, education and other items.

Hypothesis tests show that OLS estimates suffer from endogeneity. To address the potential endogeneity of household expenditure, we use an IV approach employing log income, its square and household landowning status as instruments. Due to their imprecision, our IV estimates are neither consistent with our semi-parametric analysis nor statistically significant. To have a more precise estimate, we utilise the control function approach to control for the endogeneity of both household expenditure and its square. We compare these models against corresponding models which only control for the endogeneity of the household expenditure. The latter model has better fit for all categories. Therefore, we finally estimate models with only controlling for the endogeneity of the log expenditure. In these estimates, Engel curves for all the categories except medical expenses are quadratic.

The food Engel curve for Bangladesh seems to be quadratic. This is different from the food Engel curve typically found for developed countries, which appears to be Working-Leser. Our analysis thus provides additional evidence to support the hypothesis that quadratic Engel curves are a feature of developing countries. Importantly, in case of a negative income shock, misspecified models for the Engel curve underestimate the expenditure variability of low-income households. Models accounting for the shape of the Engel curve may thus provide a more appropriate policy guideline and better forecasts.

Tables and Figures

Table 1: Family Composition

	Number of kids					
Number of adults	0	1	2	3	4	Total
1	115	193	248	156	53	765
2	723	1,515	1,998	1,173	468	5,877
3	368	622	713	413	183	2,299
4	181	331	327	172	83	1,094
Total	1,387	2,661	3,286	1,914	787	10,035

Table 2: Summary Statistics: Share of Expenditures

	Food	Clothing	Transport	Education	Medical	Other
Mean	0.61	0.07	0.05	0.04	0.03	0.20
SD	0.13	0.03	0.05	0.06	0.05	0.10
Min	0.08	0.00	0.00	0.00	0.00	0.00
Max	1.00	0.40	0.81	0.63	0.62	0.81
Percentiles of Expenditure Share						
01	0.27	0.02	0.00	0.00	0.00	0.06
05	0.36	0.03	0.00	0.00	0.00	0.09
10	0.42	0.04	0.01	0.00	0.00	0.10
25	0.52	0.05	0.02	0.00	0.01	0.13
50	0.62	0.07	0.04	0.01	0.02	0.18
75	0.70	0.09	0.07	0.06	0.04	0.24
90	0.76	0.11	0.10	0.11	0.08	0.32
95	0.79	0.13	0.14	0.15	0.12	0.39
99	0.84	0.17	0.25	0.28	0.24	0.54

Number of Observations - 10,035

Table 3: Mean of Expenditure Shares at Different Income Quintiles

	Quintile-1	Quintile-2	Quintile-3	Quintile-4	Quintile-5
Food	0.66	0.65	0.63	0.59	0.50
Clothing	0.07	0.07	0.07	0.07	0.06
Transport	0.04	0.04	0.05	0.06	0.07
Education	0.01	0.02	0.03	0.05	0.08
Medical	0.03	0.03	0.03	0.04	0.04
Other	0.18	0.18	0.18	0.20	0.25

Table 4: Summary Statistics: Independent Variables

	Mean	SD
Household Finances		
HH Consumption Expenditure (Monthly)	9,454	5,079
Household Income (Monthly)	10,760	18,213
Per capita Expenditure	2,404	1,411
Per capita Income	2,812	4,938
Demographics		
HH Head's Age	44.62	13.52
Family Size	4.18	1.34
Number of Adults in the HH	2.37	0.78
Number of Kids in the HH	1.81	1.13
Education		
HH Head has no Education	0.53	0.50
HH Head's Spouse has no Education	0.53	0.50
HH Head has Primary Education	0.16	0.36
HH Head's Spouse has Primary Education	0.17	0.37
HH Head has Secondary Education	0.22	0.41
HH Head's Spouse has Secondary Education	0.26	0.44
HH Head has Higher Secondary Education	0.07	0.26
HH Head's Spouse has Higher Secondary Education	0.04	0.19
HH Head has Graduate degree	0.02	0.14
HH Head's Spouse has Graduate degree	0.01	0.08
Employment		
Proportion of Working Men	0.27	0.16
Proportion of Working Women	0.04	0.12
Other		
Urban	0.36	0.48
Lean Season	0.17	0.37

Number of Observations - 10,035

Note: We also use information on regions and land ownership – variables closely related to the economic status of the household.

Table 5: Estimated Equivalence Scales

	Kids				
Adult	0	1	2	3	4
1	1.00	1.12	1.24	1.36	1.47
2	1.67	1.77	1.88	1.98	2.07
3	2.25	2.35	2.44	2.53	2.62
4	2.79	2.88	2.96	3.05	3.13

Table 6: Hardle and Mammen (1993) Test Results

Expenditure Shares	p-value
Food	0.24
Clothing	0.12
Transport	0.06
Medical	0.39
Education	0.01
Other	0.89

Note: Ho: The SP fit can be approximated by a quadratic fit.

Table 7: OLS Estimate of Expenditure Shares

	Food	Clothing	Transport	Medical	Education	Other
Log of Equivalised	0.558***	0.056***	0.108***	-0.019	-0.097***	-0.607***
HH Expenditure	(0.074)	(0.018)	(0.034)	(0.031)	(0.035)	(0.069)
Squared Log of	-0.040***	-0.004***	-0.005**	0.002	0.008***	0.040***
Equivalised HH Expenditure	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)
Adjusted R^2	0.379	0.114	0.102	0.042	0.242	0.226
F	228.0	46.0	36.8	12.3	103.8	85.6
N	10,035	10,035	10,035	10,035	10,035	10,035

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Test of Exogeneity & Test of Overidentifying Restrictions

	Test of Exogeneity	Test of Over-identifying Restrictions
Expenditure Shares	(p-values)	(p-values)
Food	0.000	0.065
Clothing	0.000	0.102
Transport	0.002	0.337
Medical	0.000	0.057
Education	0.410	0.380
Other	0.000	0.265

Note: *Test of Exogeneity:* Wooldridge's (1995) robust score test and a robust regression-based test are reported. If the test statistic is significant, the variables being tested must be treated as endogenous. *Tests of Overidentifying Restrictions:* Basmann's (1960) chi-squared tests are reported, as is Wooldridge's (1995) robust score test. A statistically significant test statistic always indicates that the instruments may not be valid.

Table 9: IV Estimate of Expenditure Shares

	Food	Clothing	Transport	Medical	Education	Other
Log of Equivalised	0.068	0.320**	-0.619	-0.067	0.225	0.072
HH Expenditure	(0.639)	(0.161)	(0.439)	(0.291)	(0.299)	(0.504)
Squared Log of	-0.013	-0.019**	0.038	0.004	-0.012	0.001
Equivalised HH Expenditure	(0.038)	(0.010)	(0.026)	(0.017)	(0.018)	(0.030)
Adjusted R^2	0.363	0.045	0.042	0.034	0.233	0.200
N	10,035	10,035	10,035	10,035	10,035	10,035

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 10: OLS Estimate of Expenditure Shares-CF Approach

	Food	Clothing	Transport	Medical	Education	Other
Log of Equivalised	0.484***	0.088***	0.126***	-0.042	-0.102***	-0.555***
Household Expenditure	(0.075)	(0.018)	(0.035)	(0.032)	(0.036)	(0.071)
Squared log of	-0.038***	-0.005***	-0.006***	0.003	0.008***	0.038***
Equivalised HH Expenditure	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)	(0.004)
Residual	0.045***	-0.020***	-0.011***	0.014***	0.003	-0.032***
	(0.007)	(0.002)	(0.003)	(0.003)	(0.004)	(0.006)
Observations	10035	10035	10035	10035	10035	10035
Adjusted R^2	0.382	0.126	0.103	0.044	0.242	0.229
F	223.290	47.806	36.118	12.306	100.056	84.358

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

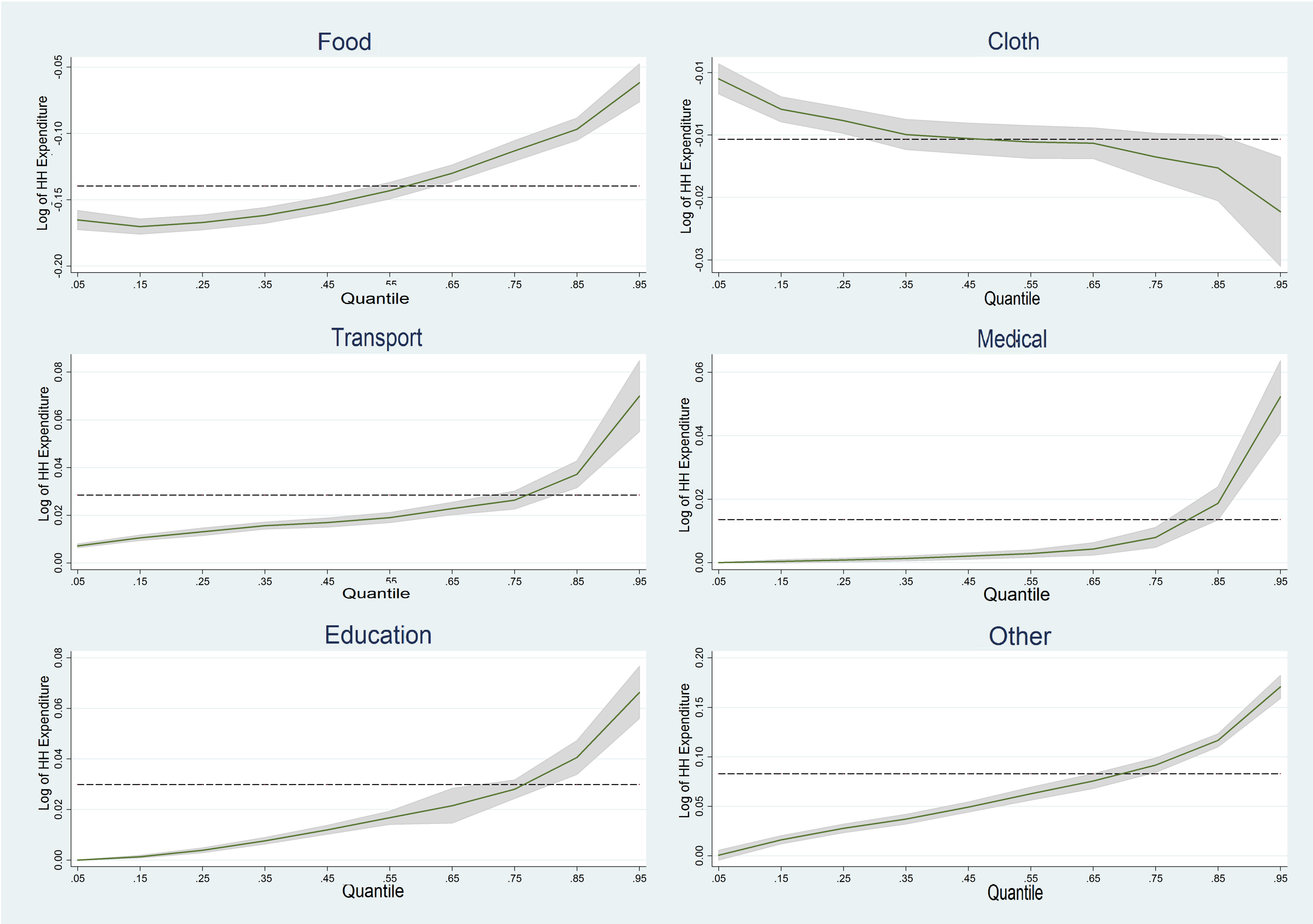

Figure 1: Plot of Coefficients at Different Quantiles

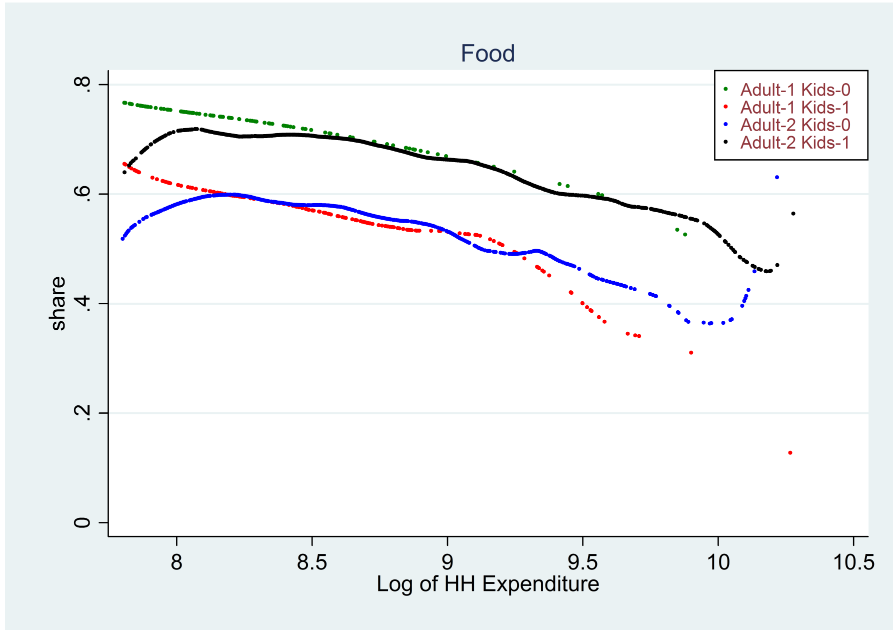

Figure 2: Semi-parametric Estimate of Food share for Different Family Types

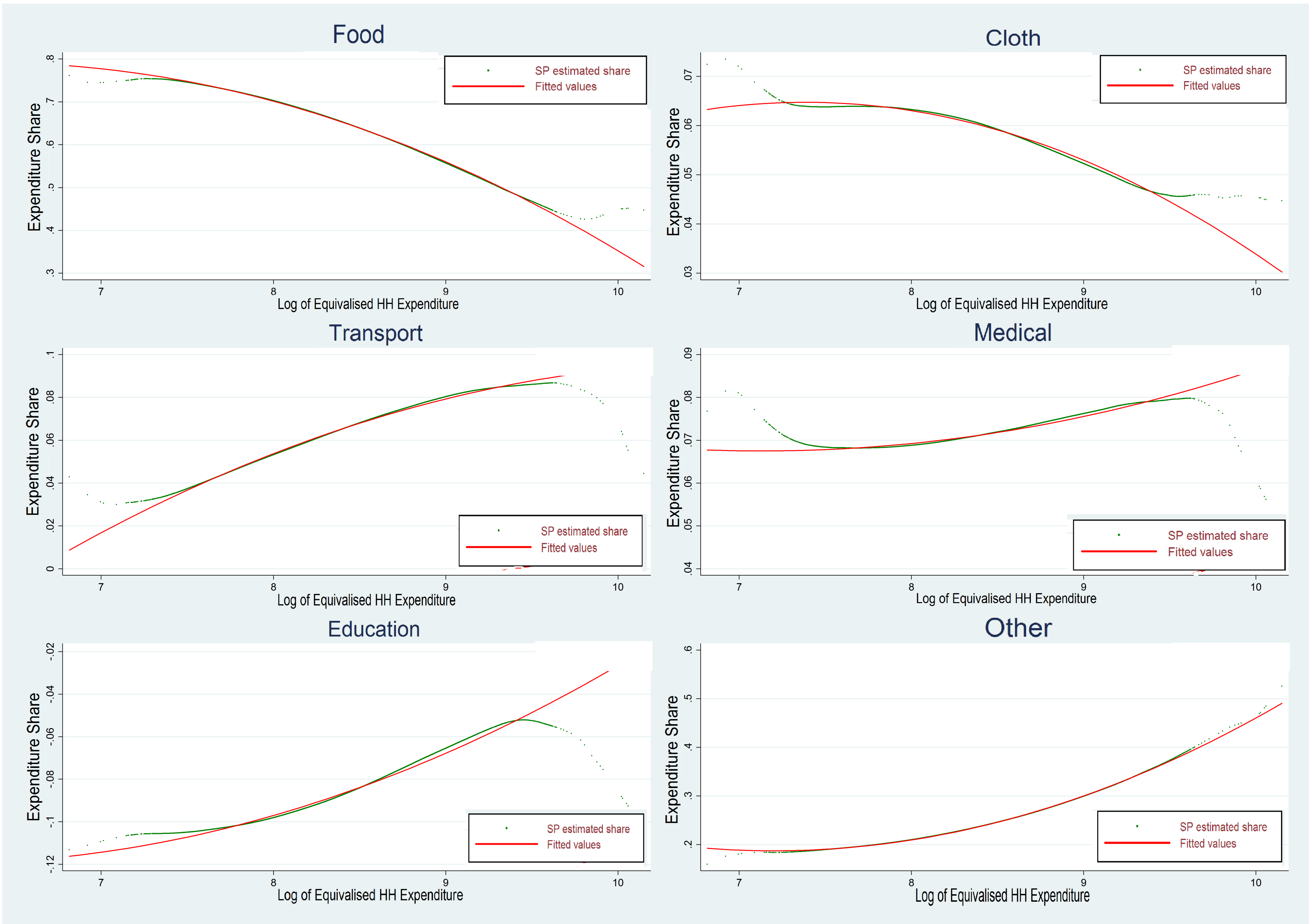

Figure 3: Semi-parametric Estimate of Base Independent Engel Curve and Quadratic Fit

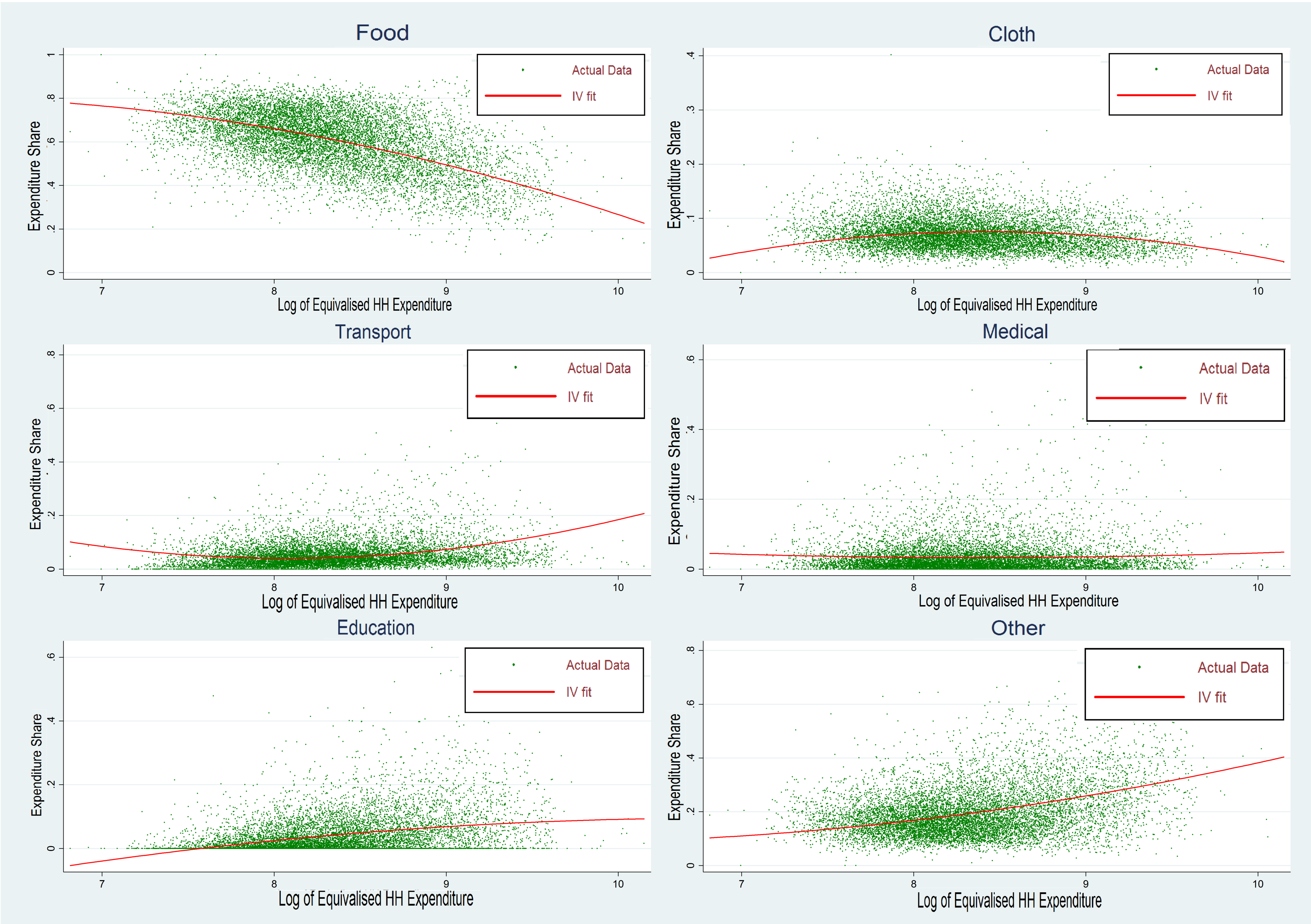

Figure 4: Actual Data and IV Fitted Quadratic Engel Curve

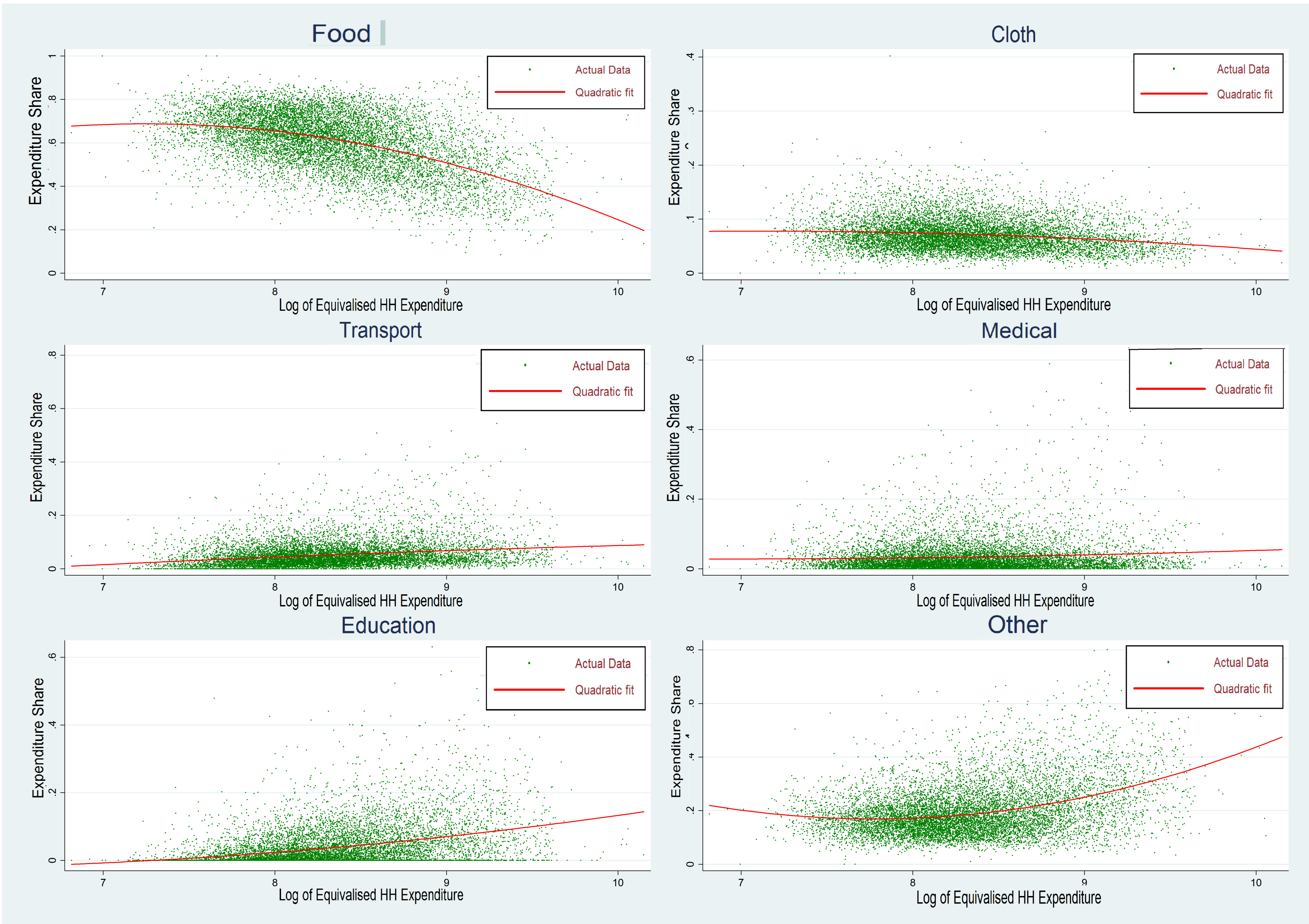

Figure 5: Actual Data and OLS Fitted Quadratic Engel Curve-CF Approach

A Appendix

A.1 Semi-parametric Model Estimation Technique

Semi-parametric estimation technique in this paper follows Robinson (1988). First, we predict the dependent and all the independent variables non-parametrically from log expenditure. Second, for the dependent and all the independent variables, we obtain the difference between actual and predicted value of each variable. Third, we use OLS to estimate the coefficients of the independent variables, by regressing the differenced dependent variable on the differenced independent variables, which enter parametrically into the model. We use the estimated coefficients to estimate the impact of these variables on the dependent variables. Now we subtract these estimated values (impact) from the dependent variable, so that we are only left with the impact of log expenditure on the dependent variable. Finally, we again run non-parametric regression of the impact free variable on log expenditure.

With $\bar{y}$ representing equivalised expenditure, our semi-parametric model is[23]

$$w_i = F(log\bar{y}) + V\lambda_i + \upsilon \tag{11}$$

If household expenditure is uncorrelated with error, conditional expectation of (11) gives,

$$E[w_i|log\bar{y}] = F(log\bar{y}) + E[V|log\bar{y}]\lambda_i \tag{12}$$

Estimates of the conditional moments can be found through bivariate non-parametric local linear regression. Subtracting (12) from (11) gives,

$$w_i - E[w_i|log\bar{y}] = (V - E[V|log\bar{y}])\lambda_i + \upsilon \tag{13}$$

The vector λ_i can be estimated by OLS using (13). We can use these estimates along with the estimated conditional moments in (12) to have an estimate of $F(log\bar{y})$,

[23]This section heavily borrows from Breunig and McKibbin (2012).

$$\widehat{F(log\bar{y})} = E[\widehat{w_i|log\bar{y}}] - E[\widehat{V|log\bar{y}}]\hat{\lambda}_i \tag{14}$$

In modelling Engel curve, household expenditure is expected to suffer from endogeneity. To deal with that, we predict the residuals from non-parametric estimation of log expenditure on log income. Then we use those residuals as an additional covariate while estimating equation(13) by OLS. Such procedure generates consistent estimates of the covariates while the significance of the residuals may also indicate the presence of endogeneity.

References

Bangladesh Bureau of Statistics. Report of the household income and expenditure survey - 2010. Report, Bangladesh Bureau of Statistics (BBS), Government of the People's Republic of Bangladesh, April 2012.

J. Banks, R. Blundell, and A. Lewbel. Quadratic engel curves and consumer demand. *The Review of Economics and Statistics*, 79(4):pp. 527–539, 1997. ISSN 00346535. URL http://www.jstor.org/stable/2951405.

S. Bhalotra and C. Attfield. Intrahousehold resource allocation in rural pakistan: A semiparametric analysis. *Journal of Applied Econometrics*, 13(5):pp. 463–480, 1998. ISSN 08837252. URL http://www.jstor.org/stable/223220.

H. Bierens and H. Pott-Buter. Specification of household engel curves by nonparametric regression. *Econometric Reviews*, 9(2):123–184, 1990. URL http://ideas.repec.org/a/taf/emetrv/v9y1990i2p123-184.html.

C. Blackorby and D. Donaldson. Adult-equivalence scales and the economic implementation of interpersonal comparisons of well-being. *Social Choice and Welfare*, 10:335–361, 1993. ISSN 0176-1714. URL http://dx.doi.org/10.1007/BF00182510. 10.1007/BF00182510.

R. Blundell and A. Duncan. Kernel regression in empirical microeconomics. *The Journal of Human Resources*, 33(1):pp. 62–87, 1998. ISSN 0022166X. URL http://www.jstor.org/stable/146315.

R. Blundell, A. Duncan, and K. Pendakur. Semiparametric estimation and consumer demand. *Journal of Applied Econometrics*, 13(5):435–461, 1998. ISSN 10991255. URL http://dx.doi.org/10.1002/(SICI)1099-1255(1998090)13:5<435::AID-JAE506>3.0.CO;2-K.

R. Blundell, X. Chen, and D. Kristensen. Semi-nonparametric iv estimation of shape-invariant engel curves. *Econometrica*, 75(6):pp. 1613–1669, 2007. ISSN 00129682. URL `http://www.jstor.org/stable/4502045`.

R. V. Breunig and D. Cobb-Clark. Understanding the factors associated with financial stress in australian households. *Australian Social Policy 2005*, 2005. URL `http://www.ssrn.com/paper=1728996`.

R. V. Breunig and R. J. McKibbin. Income pooling between australian young adults and their parents. *LABOUR*, 26(2):235–265, 2012. ISSN 1467-9914. doi: 10.1111/j.1467-9914.2011.00541.x. URL `http://dx.doi.org/10.1111/j.1467-9914.2011.00541.x`.

A. C. Cameron and P. K. Trivedi. *Microeconometrics: Methods and Applications*. Cambridge University Press, New York, USA, 1st edition, 2005. ISBN 0521848059.

A. Deaton and J. Muellbauer. *Economics and Consumer Behaviour*. Cambridge University Press, Cambridge, USA, 1st edition, 1980. ISBN 0521296765.

J. Engel and A. Kneip. Recent approaches to estimating engel curves. *Journal of Economics*, 63:187–212, 1996. ISSN 0931-8658. URL `http://dx.doi.org/10.1007/BF01258672`.

P. L. Gozalo. Nonparametric bootstrap analysis with applications to demographic effects in demand functions. *Journal of Econometrics*, 81(2):357 – 393, 1997. ISSN 0304-4076. doi: 10.1016/S0304-4076(97)86571-2. URL `http://www.sciencedirect.com/science/article/pii/S0304407697865712`.

D. Grigg. Food expenditure and economic development. *GeoJournal*, 33:377–382, 1994. ISSN 0343-2521. URL `http://dx.doi.org/10.1007/BF00806419`. 10.1007/BF00806419.

W. Hardle and E. Mammen. Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, 21(4):pp. 1926–1947, 1993. ISSN 00905364. URL `http://www.jstor.org/stable/2242322`.

W. Hardle and J. S. Marron. Optimal bandwidth selection in nonparametric regression function estimation. *The Annals of Statistics*, 13(4):pp. 1465–1481, 1985. ISSN 00905364. URL http://www.jstor.org/stable/2241365.

C. A. Knight and D. D. Ackerly. Variation in nuclear dna content across environmental gradients:a quantile regression analysis. *Ecology Letters*, 5(1):66–76, 2002. ISSN 1461-0248. doi: 10.1046/j.1461-0248.2002.00283.x. URL http://dx.doi.org/10.1046/j.1461-0248.2002.00283.x.

R. Koenker and K. F. Hallock. Quantile regression. *The Journal of Economic Perspectives*, 15(4):pp. 143–156, 2001. ISSN 08953309. URL http://www.jstor.org/stable/2696522.

A. Lewbel. Shape-invariant demand functions. *The Review of Economics and Statistics*, 92(3):549–556, 09 2010. URL http://ideas.repec.org/a/tpr/restat/v92y2010i3p549-556.html.

N. Liviatan. Errors in variables and engel curve analysis. *Econometrica*, 29(3):pp. 336–362, 1961. ISSN 00129682. URL http://www.jstor.org/stable/1909636.

Ministry of Finance. Bangladesh economic review. Technical report, June 2012. URL http://www.www.mof.gov.bd.

N. Y. Naik and M. J. Moore. Habit formation and intertemporal substitution in individual food consumption. *The Review of Economics and Statistics*, 78(2):321–28, May 1996. URL http://ideas.repec.org/a/tpr/restat/v78y1996i2p321-28.html.

W. K. Newey, J. L. Powell, and F. Vella. Nonparametric estimation of triangular simultaneous equations models. *Econometrica*, 67(3):pp. 565–603, 1999. ISSN 00129682. URL http://www.jstor.org/stable/2999545.

K. Pendakur. Semiparametric estimates and tests of base-independent equivalencescales. *Journal of Econometrics*, 88(1):1 – 40, 1999. ISSN 0304-4076. doi: 10.1016/

S0304-4076(98)00020-7. URL http://www.sciencedirect.com/science/article/pii/S0304407698000207.

P. M. Robinson. Root-n-consistent semiparametric regression. *Econometrica*, 56(4):pp. 931–954, 1988. ISSN 00129682. URL http://www.jstor.org/stable/1912705.

R. Summers. A note on least squares bias in household expenditure analysis. Cowles Foundation Discussion Papers 29, Cowles Foundation for Research in Economics, Yale University, 1957. URL http://ideas.repec.org/p/cwl/cwldpp/29.html.

V. Verardi and N. Debarsy. Robinson's squareroot-of-n-consistent semiparametric regression estimator in stata. Working Papers 1110, University of Namur, Department of Economics, Mar. 2011. URL http://ideas.repec.org/p/nam/wpaper/1110.html.

J. M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*, volume 1 of *MIT Press Books*. The MIT Press, 2010. URL http://ideas.repec.org/b/mtp/titles/0262232588.html.

H. Working. Statistical laws of family expenditure. *Journal of the American Statistical Association*, 38(221):pp. 43–56, 1943. ISSN 01621459. URL http://www.jstor.org/stable/2279311.

A. Yatchew. Nonparametric regression techniques in economics. *Journal of Economic Literature*, 36(2):pp. 669–721, 1998. ISSN 00220515. URL http://www.jstor.org/stable/2565120.

A. Yatchew. *Semiparametric Regression for the Applied Econometrician*. Number 9780521812832 in Cambridge Books. Cambridge University Press, 2003. URL http://ideas.repec.org/b/cup/cbooks/9780521812832.html.

A. Yatchew, Y. Sun, and C. Deri. Efficient estimation of semiparametric equivalence scales with evidence from south africa. *working paper, University of Toronto,* *Dept. of Economics*, 2000. URL http://www.economics.utoronto.ca/yatchew/EquivalenceScalesWorkingPaper.pdf.

A. Yatchew, Y. Sun, and C. Deri. Efficient estimation of semiparametric equivalence scales with evidence from south africa. *Journal of Business & Economic Statistics*, 21(2):247–257, 2003. doi: 10.1198/073500103288618936. URL http://www.tandfonline.com/doi/abs/10.1198/073500103288618936.