Discussion Paper No. 97-06

Estimation of Household Demand Systems Using Unit Value Data

Ian Crawford, François Laisney and Ian Preston

ZEW

Zentrum für Europäische Wirtschaftsforschung GmbH

Centre for European Economic Research

Discussion Paper No. 97-06

Estimation of Household Demand Systems Using Unit Value Data

Ian Crawford, François Laisney and Ian Preston

Estimation of Household Demand Systems Using Unit Value Data

Ian Crawford[1], François Laisney[2] and Ian Preston[3]

ZEW

April 1997

Abstract

Many budget surveys present the interesting features that for a wide range of goods they contain quantity information along with expenditure information, and that the geographical location of households is fairly precise. We take advantage of these features to develop a method for estimation of price reactions using unit value data which exploits the implicit links between quantity and unit value choices. This allows us to combine appealing Engel curve specifications with a model of quality choice in a way which is consistent with demand theory. The method is applied to Czech data.

Key Words: Consumer demand, unit values, quality.

JEL Classification: D11, D12

Acknowledgements: We are grateful for comments and advice from Richard Blundell, Fiona Coulter, Angus Deaton, Philippe De Vreyer, Bas Donkers, Vera Kameničkova, Colin Lawson, Arthur Lewbel, Edmond Malinvaud, Costas Meghir, Daniel Munich, Jean-Marc Robin, David Ulph, Jiří Večernik, Frank Windmeijer and participants at conferences in Istanbul, Leuven, Mannheim, Prague and various seminars. This research was undertaken with support from two programmes of the European Commission: PHARE-ACE (project "Tax Policy During Economic Transition") and HCM (project "Microeconometrics of Public Policy Issues"). All errors remain our own.

1 Institute for Fiscal Studies, London. ian_crawford@ifs.org.uk

2 Université Louis Pasteur, Strasbourg, and ZEW, Mannheim. laisney@zew.de

3 Institute for Fiscal Studies and University College London. ian_preston@ifs.org.uk

Summary

In empirical demand analysis based on household data it is usually difficult to estimate price reactions, because of insufficient variation. However, many budget surveys — especially in Eastern European countries and in developing countries — present the following interesting feature: on top of the typical information on expenditures, they also contain information on quantities purchased. This allows the computation of household-specific unit values, as ratios of expenditure over quantity. The variation in these unit values over the sample for each good will result from both geographical variation in prices and from household choices regarding the quality of the good purchased, i.e. the composition of the corresponding aggregate. A typical example is meat, where the average per kilo expenditure of a household depends on the mix of beef, pork, etc. that the household consumes. In this respect, the unit values depend on household choices, and it turns out that important links exist between quantity and unit value choices. With information on the geographical location of households it is possible to separate the price and the quality components of unit values, and to use these links in the estimation of price reactions. Here we develop a new approach which has the advantage over currently used alternatives of allowing us to combine appealing budget share specifications with a model of quality choice in a way which is fully consistent with demand theory. In an application using data from the Czech Family Budget Survey for 1991 and 1992, we consider six categories of food, plus clothing and footwear and estimate a demand system conditionally on expenditures on several other goods, durable ownership and labour market status. We find that both taking quality into account and conditioning in this way are important, even though the quality elasticities appear small. The woman's participation to the labour market has a significant impact on four budget shares: negative on meat and starches, and positive on clothing and on footwear. It has a positive and significant impact on all unit values, except for dairy and vegetables/fruit. Several durables have contrasted effects on quantity and quality. Meat, alcohol and clothing appear as luxuries, vegetables/fruit and dairy products as necessities. All significant own price elasticities are negative. Several cross price effects are significant. Thus dairy products and vegetables/fruit are complements to meat, but substitutes for each other; clothing and footwear also are complementary.

1. Introduction

A main difficulty in the estimation of demand systems using household data concerns the precise estimation of price reactions. While requiring care in the treatment of endogeneity, income effects are more easily estimated. But unless, say, one is prepared to make strong assumptions on functional form which result in a connection between price and income effects, but which, if wrong, produce important biases in the estimation of price elasticities (see e.g. Deaton and Muellbauer, 1980), price effects are difficult to capture. The reason is that whereas data on households normally exhibit considerable variation in expenditures, this is not typically the case for prices. Very often information about geographical variation in prices or variation over time within the period covered by one cross-section is lacking, so that prices are assumed uniform over all households of the same cross-section. Indeed most studies based on the Family Expenditure Survey for instance — a long series of cross-sections of UK households — have relied solely on year-to-year variation of prices under that assumption (for a recent example, see Banks, Blundell and Lewbel, 1996). In the absence of such long series, researchers have often resorted to combining a small number of cross-sections with aggregate time series data, the idea being basically to identify the income effects from the cross-sectional data and the price effects from the aggregate data (examples of studies relying on that strategy are Stone, 1954, Jorgenson, Lau and Stoker, 1982, and Nichèle and Robin, 1995).

Data sets which contain information, not only on expenditures, but also on *quantities* consumed, offer interesting possibilities: this allows the computation of individual *unit values* for the spending of each household on any good for which this is true. It might be thought possible to model demand for these goods treating these "unit values" as prices. These would appear much more attractive for estimation purposes than aggregate prices, which are just averages that no household actually pays. Yet, since the "goods" are invariably subject to some degree of aggregation, it is undoubtedly true that much of the variation in unit values will actually result from household choice regarding the nature of the goods purchased.

Deaton (1987, 1988, 1990, 1995) has developed a way of modelling price reactions jointly with choice of unit values in data of this type, under assumptions about fixity of underlying relative prices within spatially defined areas. We develop an alternative though strongly related approach which exploits the implicit links between quantity and unit value choices. This allows us to combine appealing Engel curve specifications with a model of quality choice in a way which is consistent with demand theory.

We use the Czech Family Budget Survey which has the feature that the geographical location of households is fairly precise. The preference specification used is the Almost Ideal Demand (AID) system and the eight goods categories retained are six categories of food, plus clothing and footwear. In order to avoid arbitrary separability assumptions, the demand system is estimated conditionally on expenditures on several other good categories, on durable ownership and on labour market status.

The results are encouraging, and our approach has subsequently been applied with success in the context of the estimation of heterogeneous labour demand functions by De Vreyer (1996).

Section 2 discusses relevant points from demand theory. Section 3 outlines a three stage estimation methodology. Section 4 describes the Czech data used and Section 5 presents illustrative results.

2. Demand and unit values

We start with a development of Deaton's approach to modelling the determination of unit values. For the purpose of empirical investigation, goods are taken to be organised into m groups such as meat, fish, clothes and so on. Consumption within a group G is a vector of quantities q_G. A group quantity index Q_G is defined as

$$Q_G \equiv k_G . q_G \ , \tag{2.1}$$

where k_G is a vector of aggregating units typically chosen by the data collector (like weight for meat, or pairs for shoes, but which could also be a characteristic like calories if they were observed; k_G could even be a reference price vector). Group spending is

$$x_G \equiv p_G \cdot q_G \ ,$$

where p_G is the vector of prices.

We follow Deaton in making two central assumptions. Firstly, we assume that relative prices *within* each group are fixed, so that $p_G = \pi_G p_G^0$, where π_G is defined as a scalar (Paasche) linear homogeneous price *level* for the group (for instance, the price of meat relative to other groups), and p_G^0 is a vector representing the fixed within-group relative price structure (for instance, the relative prices of different types and qualities of meat). This assumption will allow us to treat group G as a Hicks aggregate, so that x_G will be a function of the vector π of group price levels (generally, omission of a G subscript for a group variable will denote the vector of values for all groups) and total spending $X = \sum_G x_G$.[1] Secondly, again following Deaton, the price vector π is assumed constant within identifiable regional clusters of households, so that we will use the notation π^c for clusters $c = 1, \ldots, C$. This is the central identifying assumption of Deaton's approach.

For each household we observe a *unit value* for each group of goods:

$$V_G \equiv \frac{x_G}{Q_G} \ , \tag{2.2}$$

and given the definitions above, this leads naturally to a new concept: since

$$\frac{x_G}{Q_G} = \frac{\pi_G p_G^0 \cdot q_G}{k_G \cdot q_G} \ , \tag{2.3}$$

we have

$$V_G \equiv \pi_G \xi_G \ , \tag{2.4}$$

where ξ_G, defined as

[1] The assumption appears very strong, but Lewbel (1996) shows that this type of aggregation will be possible under the much weaker assumption of stochastic independence between π_G and the vector of relative prices p_G/π_G. This will be the case, at least approximately, if the relative prices are stationary over time, whereas π_G is not. We will come back to this assumption when discussing the stochastic structure of the econometric specification, as it will turn out to be an important identifying assumption.

$$\xi_G \equiv p_G^0 \cdot q_G / Q_G \ , \tag{2.5}$$

is an index of *expensiveness* for group G at reference prices.[2] It is interpreted by Deaton as an indicator of *quality* though such an interpretation is unimportant to the suggested procedure for derivation of the correct price responses for quantities.

Given the assumed fixity of p_G^0, the variables Q_G, x_G, ξ_G, V_G, as well as related variables such as budget shares,

$$w_G \equiv x_G / X \ , \tag{2.6}$$

will all be functions of X and π.

It is important to note that there are restrictions between the unit value and budget share equations which should lead one to be cautious before proceeding with independent specification choices for $w_G(X, \pi)$ and $V_G(X, \pi)$. Assuming weak separability of preferences in the partition corresponding to groups $1, \ldots, G, \ldots, m$, (as Deaton does), and using Hicks aggregation and homogeneity, we can write

$$\begin{aligned} q_G &= \widetilde{f}_G(x_G, p_G) \\ &= \widetilde{f}_G(x_G/\pi_G, p_G^0) \\ &= f_G(x_G/\pi_G) \\ &= f_G(\xi_G Q_G) \end{aligned} \tag{2.7}$$

(suppressing dependence on p_G^0). Thus, since both Q_G and ξ_G depend only on q_G,

$$\begin{aligned} Q_G &= H_G(\xi_G Q_G) \\ \xi_G &= h_G(\xi_G Q_G) \ . \end{aligned} \tag{2.8}$$

These equations, which make clear the cross-equation restrictions on the functional forms of quantity and unit value equations, are central to our treatment. Though never explicitly stated, (2.8) is implicitly used by Deaton since it underlies the equation

[2] If k_G were selected equal to p_G^0, then ξ_G would be identically equal to 1. This choice is not open to us because we do *not* observe the relative prices.

$$\frac{\partial \ln \xi_G / \partial \ln \pi_H}{\partial \ln \xi_G / \partial \ln X} = \frac{\partial \ln Q_G / \partial \ln \pi_H}{\partial \ln Q_G / \partial \ln X} , \tag{2.9}$$

used to derive price elasticities at the second stage of his estimation procedure.

If both the quantity and unit value relationships are specified to be double logarithmic, as, for example, in the studies by Deaton (1987, 1988), then this specification is compatible with (2.8).[3] However, difficulties arise if the method is applied with other functional forms. If, for instance, an AID type budget share equation — with w linear in $\ln X$ and $\ln\pi$ — is adopted while the log unit value is also specified linearly in the same variables, this is not compatible with (2.8) (except under extremely strong restrictions - see Appendix C for details).[4] From (2.6) and (2.8) it follows that one would need to estimate a consistent system

$$\begin{aligned} w_G &= w_G(X,\pi) \ , \\ \ln V_G &= \ln \pi_G + \ln h_G \left[\frac{X}{\pi_G} w_G(X,\pi) \right] \ . \end{aligned} \tag{2.10}$$

A simple linear specification in $\ln X$ and $\ln\pi$ for the share equation therefore requires a unit value equation which will be non-linear in these variables. The problem here is that, once the quantity or budget share relationship is specified, (2.8) imposes too many cross-equation restrictions to permit also an unrestricted dependence of unit values on X and π.

Our suggestion is to specify the quantity or budget share relationship, $w_G(X,\pi)$, and then to derive a relationship between V_G and Q_G from an independent specification of (2.8) (since the form of h_G is unrestricted). To be more specific, if we posit a share equation such as

$$w_G = \alpha_G + \sum \delta_{GH} \ln \pi_H + \beta_G \ln X , \tag{2.11}$$

[3] Yet there are very strong restrictions on the coefficients: if $\ln Q_G = \alpha_G + \beta_G \ln X + \sum_H \gamma_{GH} \ln \pi_H$ and $\ln \xi_G = a_G + b_G \ln X + \sum_H c_{GH} \ln \pi_H$ then by (2.9) $\beta_G / b_G = \gamma_{GH} / c_{GH}$ for all H.

[4] This specification has been adopted in Deaton (1990), Deaton and Grimard (1991) and Ayadi, Baccouche, Goaied and Matoussi (1995). In none of these papers is the incompatibility explicitly recognised, although Deaton (1995) suggests that it might be appropriate to use (2.9) at mean sample assuming constancy of elasticities as a reasonable approximation to the truth.

then the functional form

$$\ln V_G = a_G + \sum d_{GH} \ln \pi_H + b_G \ln X$$

is not allowed for the log unit values. Yet the specification of

$$h_G(\xi_G Q_G) = (\xi_G Q_G)^{b_G/(1+b_G)} \exp\left[a_G/(1+b_G)\right]$$

leads to a simple form for the latter:

$$\begin{aligned} \ln \xi_G &= a_G + b_G \ln Q_G \\ \Longrightarrow \ln V_G &= a_G + b_G \ln Q_G + \ln \pi_G. \end{aligned} \tag{2.12}$$

3. Econometric considerations

From (2.6), and choosing for our specification the approximate AID model with a log-linear approximation to the log price index, the share equation for good i, demanded by household h in cluster c, is given by

$$w_G^h = \alpha_{0G} + Z^h \alpha_G + \sum_H \gamma_{GH} \ln \pi_H^c + \beta_G \ln \check{x}^h + u_G^h \tag{3.1}$$

where $\ln \check{x}^h \equiv \ln X^h - \ln P^c \equiv \ln X^h - \sum_H \lambda_H \ln \pi_H^c$, $\check{x}^h$ is deflated expenditure and P^c is a cluster price index for suitably chosen λ.[5] This leads to the equation

$$w_G^h = \alpha_{0G} + Z^h \alpha_G + \sum_H \delta_{GH} \ln \pi_H^c + \beta_G \ln X^h + u_G^h, \tag{3.2}$$

with $\delta_{GH} = \gamma_{GH} - \beta_G \lambda_H$. Vector Z^h includes socio-demographic characteristics as well as further conditioning variables, mentioned in the introduction, and which will be described in detail in the next section. Several of these are potentially endogenous and will be instrumented.

For (2.12) we assume that the unit value equation is of the form

$$\ln V_G^h = a_{0G} + Z^h a_G + \ln \pi_G^c + b_G \ln Q_G^h + v_G^h . \tag{3.3}$$

We assume independence between observations. This may appear unduly restrictive, as it rules out the presence of cluster effects. But firstly, we have to rule

[5] It may seem overrestrictive to impose constancy of the weights λ across clusters. However, relaxing that assumption, for instance in order to specify P^c as a Stone price index for cluster c, with λ^c the vector of average budget shares in cluster c, would lead to cluster-specific coefficients γ or δ.

out the simultaneous appearance of cluster effects in both share and unit value equations, as this would preclude the identification of the price effects. Secondly, allowing cluster effects in the share equation only would not change anything in the sequel, provided these effects were independent of π. This is where Lewbel's assumption is helpful again: allowing the relative prices p_G^0 to vary across clusters — and thus become p_G^{0c} — introduces a cluster effect which depends on the latter; assuming independence between p_G^{0c} and π^c makes this cluster effect innocuous.[6] And thirdly, postulating additive errors in these equations is questionable anyway, as equation (2.10) shows.

The covariance matrix Ω of the vectors $\left(u^{h'}, v^{h'}\right)'$ is assumed constant across observations and otherwise unrestricted. This homoscedasticity assumption is less plausible for log unit values than it is for budget shares, but we reckon that it would be difficult to relax it in the quality model, as should become apparent.

A first strategy might be to estimate (3.2) replacing prices with unit values while instrumenting the latter. An approach of this type has been adopted by Pitt (1983) and Strauss (1982). The implicit assumption of such an approach is that the vector of unit values V^h is simply an error-ridden observation of the price vector π^c, with a measurement error that is independent of π^c. In the context of our quality choice model this amounts to the assumption that all parameters shown in equation (3.3) are zero, so that the quality index ξ^h does not depend from the outlay X^h and the vector π^c of price indices.[7] This points to the likely misspecification of this approach since, should this assumption fail to hold, the parameters of (3.2) will not be properly recovered.

The estimation proceeds in three stages. In the first stage we estimate for each good a share equation and a log unit value equation using within cluster

[6] Thanks to Philippe De Vreyer for having pointed this out.

[7] De Vreyer (1996) shows that a sufficient condition for this is homothetic separability of preferences.

estimation and instrumental variables in a 2SLS framework.[8] In the second stage we retrieve the price coefficients using between cluster estimation while taking account of measurement errors on the unit values. The third stage imposes the symmetry restrictions through minimum distance estimation.

3.1 First stage

Averaging (3.2) over households within the cluster c yields

$$\overline{w_G}^c = \alpha_{0G} + \overline{Z}^c \alpha_G + \sum_H \delta_{GH} \ln \pi_H^c + \beta_G \overline{\ln X}^c + \overline{u_G}^c. \tag{3.4}$$

The vector $\widehat{\alpha}_G$ and the scalar $\widehat{\beta}_G$ are recovered from within cluster estimation, i.e. the estimating equation is obtained by subtracting (3.4) from (3.2). Similarly, forming cluster means from (3.3)

$$\overline{\ln V_G}^c = a_{0G} + \overline{Z}^c a_G + \ln \pi_G^c + b_G \overline{\ln Q_G}^c + \overline{v_G}^c , \tag{3.5}$$

$\widehat{a}_G$ and $\widehat{b}_G$ are estimated by within cluster estimation.

Endogeneity issues are addressed by use of instrumental variables where appropriate, as discussed in the data section below. Several variables are instrumented by cluster means excluding the current observation. We justify this technique as follows. In the regression $y_i - \bar{y}^c = \left(X_i - \overline{X}^c\right)\beta + u_i - \bar{u}^c$, let $\check{X}_i = \frac{1}{n_c - 1} \sum_{\substack{j \in c \\ j \neq i}} X_j$ and consider the asymptotic covariance between $\check{X}_i$ and $u_i - \bar{u}^c$: we have

$$\mathrm{E}\left[\check{X}_i \left(u_i - \bar{u}^c\right)\right] = -\mathrm{E}\left[\check{X}_i \bar{u}^c\right] = -\frac{1}{n_c}\frac{1}{n_c - 1} \sum_{\substack{j \in c \\ j \neq i}} \mathrm{E}\left[X_j u_j\right] = -\frac{1}{n_c}\mathrm{E}\left[X u\right] ,$$

which goes to zero when the number of observations per cluster goes to infinity.

[8] There are two reasons here for preferring 2SLS to the more efficient 3SLS procedure. First, 3SLS risks contaminating the estimates of the share equation by a misspecification of the unit value equation (or the reverse, but we have more confidence in the validity of the share specification). Second, 2SLS estimates of the share equations will automatically satisfy adding-up restrictions, whereas this does not necessarily hold for 3SLS estimates (see e.g. Bewley, 1986).

Further note that the within-cluster technique adopted will not only sweep away the unobservable price indices from the share equations, but also any cluster-specific effect. At this stage, the independence between cluster effects and prices plays no role, but it becomes important in the next stage.

3.2 Second stage

Separating observables and unobservables in (3.4) and (3.5) yields

$$\eta_G^c \equiv \overline{w_G}^c - \overline{Z}^c \hat{\alpha}_G - \hat{\beta}_G \overline{\ln X}^c = \alpha_{0G} + \sum_H \delta_{GH} \ln \pi_H^c + \overline{u_G}^c \equiv \eta_G^{*c} + \overline{u_G}^c \quad (3.6)$$

and

$$\zeta_G^c \equiv \overline{\ln V_G}^c - \overline{Z}^c a_G - b_G \overline{\ln Q_G}^c = a_{0G} + \ln \pi_G^c + \overline{v_G}^c \equiv \zeta_G^{*c} + \overline{v_G}^c \, . \quad (3.7)$$

Only between cluster information needs to be considered here, since no information on the price responses remains to be exploited within clusters. The true relationship between η_G^{*c} and the vector ζ^{*c} with components ζ_H^{*c} is thus

$$\eta_G^{*c} = \rho_G + \sum_H \delta_{GH} \zeta_H^{*c} = \rho_G + \zeta^{*c} \delta_G$$

with $\rho_G \equiv \alpha_{0G} - \sum_H \delta_{GH} a_{0H}$, and this suggests the regression of η_G^c on ζ^c and a constant. Measurement error bias is caused by the correlation between the vectors ζ^c, $\overline{v}^c$ and possibly $\overline{u}^c$, but is easily corrected because the variance of ($\overline{u}_G^c$, $\overline{v}^c$) can be estimated as

$$\hat{V}\begin{pmatrix} \overline{u}_G^c \\ \overline{v}^c \end{pmatrix} = \frac{1}{n_c} \begin{bmatrix} \hat{\Omega}_{u_G} & \hat{\Omega}_{u_G v} \\ \hat{\Omega}_{v u_G} & \hat{\Omega}_v \end{bmatrix}$$

where each term of $\hat{\Omega}$ is obtained from the residuals of the previous stage. This is the place where the difficulty of relaxing the homoscedasticity assumption becomes manifest. It is now easily seen that under our assumptions a consistent estimator of the vector δ_G, after demeaning the η and ξ variables and scaling them by $\sqrt{n_c}$, is given by[9]

$$\hat{\delta}_G = [\sum_{c=1}^{C} n_c \zeta^c \zeta^{c\prime} - \hat{\Omega}_v]^{-1} [\sum_{c=1}^{C} n_c \eta_G^c \zeta^{c\prime} - \hat{\Omega}_{u_G v}]. \quad (3.8)$$

3.3 Third stage

Thus far we have estimated $\hat{\beta}_G$ and $\hat{\delta}_{GH}$ for all G and $H = 1, \ldots, m$, and these estimates will automatically satisfy the adding-up restrictions, $\sum_H \hat{\beta}_H = 0$ and

[9] Asymptotics here concern the thought experiment where both the number of observations in each cluster n_c and the number of clusters C go to infinity.

$\sum_G \hat{\delta}_{GH} = 0$. The parameter estimates we are interested in are the $\hat{\lambda}_G$ and $\hat{\gamma}_{GH}$ for all G and H. The latter must satisfy symmetry, $\hat{\gamma}_{GH} = \hat{\gamma}_{HG}$, and homogeneity, $\sum_H \hat{\gamma}_{GH} = 0$, which also implies the adding-up restriction. The $\hat{\lambda}$ parameters are subject to the restrictions $\hat{\lambda}_G > 0$ and $\sum_H \hat{\lambda}_H = 1$ (positive linear homogeneity of the price index). Besides these restrictions, the relationships between the parameters of interest (γ, λ) and the auxiliary parameters $\psi = (\delta, \beta)$ are the m^2 equations $\delta_{GH} = \gamma_{GH} - \beta_G \lambda_H$. Unfortunately, these restrictions are not sufficient to identify the parameters of interest, although their number at first lured us into thinking they would. Indeed, if (γ, λ) satisfy all the restrictions, so will $(\dot{\gamma}, \dot{\lambda})$, with $\dot{\gamma}_{GH}(\kappa) = \gamma_{GH} + \kappa\beta_G\beta_H$ and $\dot{\lambda}_H(\kappa) = \lambda_H + \kappa\beta_H$, for all κ such that $\dot{\lambda}_H(\kappa) > 0$ for all H, given that $\sum_H \beta_H = 0$. Thus we regretfully set $\lambda = \bar{w}$, the vector of average budget shares, with the consequence that the price index P^c now appears as a Stone price index for cluster c, with identical weights across clusters.

We estimate symmetry-restricted parameters γ by minimum distance estimation conditional on λ. Following the efficiency arguments of Kodde, Palm and Pfann (1990, Theorem 5) we minimise only over γ rather than over γ and β. Given the linearity of the restrictions, the computations boil down to GLS estimation in the parameter space. This requires an estimate $\hat{V}(\hat{\psi})$ of the variance of the unrestricted estimator and a convenient way to obtain this is to recognise that the procedure of the first two stages falls into the framework of sequential GMM outlined by Newey (1984), as already pointed out by Deaton (1990). We briefly summarise this for completeness in Appendix C.

4. Data and specification

The data used come from the Czech Family Budget Surveys for 1991 and 1992. Data were also available for the years 1989 and 1990, but since price liberalisation dates from January 1991, it seems preferable not to use the data where behaviour would almost certainly be constrained to an extent requiring explicit treatment. Households included in the sample were asked to maintain an expenditure diary for a full twelve months, recording both quantities and expenditures for certain goods. The length of the recording period has the advantage of virtually eliminating infrequency of purchase as an explanation for zero records on most main expenditure items. However the burden imposed on participants must have been arduous and we were unable to use 480 households who did not take part over the full year. The data is a panel, but the household identifier is discarded between years, necessitating considerable effort to recover and use the panel structure. Households whose circumstances change in any major way are dropped from the sample — an unfortunate feature which again diminishes the usefulness of the panel aspect and which must also affect the cross-sectional sampling properties.

We concentrate in this paper on a subsample of married couples though it is our intention to use the whole sample in later work. The wife's labour force participation is used as a conditioning variable and instrumented. The sample size obtained in pooling the two years is 4668 households. Given that the number of identifiable geographical clusters is 179, we have an average of 26 households per cluster, with a minimum of 7 and a maximum of 60.

Eight categories of goods were selected for demand estimation. The choice was constrained by the need to have both quantities and expenditures available. Goods for which expenditures alone were available were used as conditioning goods. Detailed lists of the goods in both categories, including also the exact composition of each aggregate, can be found in Appendix A. For some commodities, the survey includes "in kind" quantities and expenditures as well as bought goods. We treat all quantities together (bought or not), defining a price index for the aggregate on the basis of the unit values.[10] A difficulty with the "in kind" records is that the corresponding unit values are constant across households, indicating that the statistical office has imputed the "expenditure" on the basis of the reported quantity, by means of a national price index.

Further variables used include socio-demographic characteristics, like age and education of the household head, his occupation, the number of persons in the household and the average age of the children, and whether or not the household lives in a rural area. Variables connected with housing are an ownership dummy,

[10] It is our intention to look further into the validity of this by investigating the behaviour of households consuming both "in kind" and through the market.

an indicator for poor housing, the average space per person, and the availability of gas supply. Durable ownership is described through the number of electronic appliances (radio and TV sets, etc.) and dummies for the possession of a freezer, a telephone, a motor vehicle, a summer house or a caravan, and a garage.[11]

As already mentioned above, our choice of goods categories to model is dictated by the availability of the information required for the construction of unit values. Since we have no reason to believe that this availability — related to the survey design — is directly connected to the structure of preferences, it is not attractive to assume that the latter are separable in the corresponding partition. Rather, following Browning and Meghir (1991), we will condition the budget shares of the included goods on the expenditures on the excluded goods. Homogeneity with respect to the prices of the excluded goods will be ensured by expressing the conditioning expenditures in relative terms with respect to one of them.[12] Furthermore we will condition the budget shares for the modelled goods on durable ownership and on a variable describing the labour market status of the household. In the words of Browning and Meghir:

> "The conditional demand system will be correctly specified whether or not [labour market status] is chosen optimally. Additionally we do not need to model explicitly the budget constraint for the conditioning goods. This is particularly significant for labour supply and for durables [...:] we may study consumer demand while being agnostic about issues such as unemployment [...] while accounting for their possible influence on demand. Conditional demand functions are an economical way of relaxing separability and still maintaining the focus on the goods of interest."

Under weak separability, these conditioning variables should play no role in the demand equations, so that we have the basis for a separability test there. The compatibility between this conditional approach and the "quality" model described above is ensured by the fact that the conditional cost function is amenable to Hicks aggregation. It will be important to remember that elasticities calculated need to be interpreted as conditional both on "total" expenditure on the

[11] Descriptive statistics on the variables used, omitted here in order to save space, are available upon request.

[12] Thanks to Arthur Lewbel for pointing this out.

modelled goods and on all these other conditioning variables.

Finally, we will also condition the unit values on variables describing durable ownership and labour market status, but not on the expenditures on other goods.

We treat as endogenous the log of total expenditure X and of quantity Q_G, the conditioning expenditures and durable ownership variables and labour market status of the wife. Instruments include the log of income (which should be correlated with $\ln X$ and $\ln Q_G$), wife's age and education and age of the youngest child (which should all be correlated with wife's participation) and cluster means of the conditioning expenditures and durable ownership excluding the current observation.

5. Estimates

We start with a description of kernel regressions for pairs of variables in order to give an impression of the sample variation in some key magnitudes of the analysis. For brevity, in Appendix D we present figures for three groups of goods only; meat, alcohol and clothing. Figures 1(a)-(c) illustrate the way in which quantities purchased and unit values vary with total expenditure and with each other. The figures show kernel regressions using the whole sample. The Engel curves for alcohol and clothing suggest that they may well be luxuries, this character being more obvious for clothing. The status of meat is ambiguous. The variation of unit values with either total spending or the quantity is clearly different for the three goods.

Estimates appear in Appendix E. In Tables 1a and 1b we report the first stage results for all goods along with their asymptotic standard errors.[13] These are the outcome of within-cluster 2SLS regressions of the type explained above: the estimating equations are obtained by subtracting (3.4) from (3.2), and (3.5) from (3.3). It is important to note that the equations presented for unit values correspond to the pure quality effects embodied in these, since the price effects

[13] Insofar as we have not attempted to identify households present in both years, these estimates should be seen as illustrative only. A particular implication of this is that the inferences drawn are based on inconsistent estimates of the variance of the estimated coefficients. An easy way out of this difficulty would be to report results separately for 1991 and 1992, which ought also to be of interest in their own right.

have been swept away.

From the unit values equations (Table 1a) we see that the assumption that the unit value is simply an error-ridden measure of price is rejected for each category, which shows that the straightforward instrumenting approach would, as suggested, be inconsistent. Clear evidence of a relationship between unit value and quantity appears only for two goods, dairy and starches, and the effect is negative.

From the share equations (Table 1b) we see that several of the conditioning goods are significant, implying decisive rejection of separability of preferences in the partition modelled goods / other goods.

The woman's participation has a significant impact on four budget shares: on meat and starches a negative impact, and a positive one on clothing and on footwear, implying also clear rejection of the separability of preferences in the partition leisure/goods. It has a positive and significant impact on all unit values, except for the two categories dairy and vegetables/fruit, and the combined effect on quantity and quality on meat and starches as opposed to clothing and footwear has a neat interpretation. Several other variables have contrasted effects on quantity and quality (see e.g. the effects of education of the household head on the quantity and quality of meat purchased), but a complete enumeration would be tedious, and the reader will be able to browse through the results without our guidance.

The suggestion that budget shares for meat, alcohol and clothing rise with total spending is confirmed: the coefficient of $\ln X$ is significantly positive. Budget shares for vegetables/fruit and for dairy products are significantly negatively affected by total spending.

It is also interesting to note the influence of the durable ownership variables: possession of a freezer, for instance, which is significant in only two of the food share equations, meat (+) and starches (-), appears to have a significant influence in five of the food unit value equations, always entering with a negative sign. To a lesser extent a similar observation can be made for motor vehicle ownership — while little evidence of an effect on budget shares is evident, there is some evidence that vehicle ownership is associated with lower unit values for some goods. The most intuitive explanations for both effects are that households thus equipped have better opportunities for purchasing in large quantities and for either taking advantage of low price opportunities or searching for them.

Finally, note that expenditure on tobacco correlates positively with the budget shares on meat and alcohol, negatively with those on dairy, starches and vegetables/fruit, while expenditure on hygiene and health go the other way round.

Second stage estimates for all equations are given in Table 2. These are estimates using the measurement error correction procedure, i.e. equation (3.8). Third stage estimates of the symmetry (and homogeneity, which follows given adding-up) restricted parameters γ_{GH} are given in Table 3. As expected, a higher proportion of the γ coefficients are significant than was the case for the δ (23/36, compared to 25/64). The table also reports the minimised value of the criterion, which provides a χ^2 test of the restrictions. We obtain a rejection at any reasonable level of significance, but remember from subsection 3.3 that the restricted estimates also embody a strong a priori restriction, namely that the deflator of total expenditure is a Stone price index which varies across clusters, but with fixed weights. Further research should investigate the precise sources of the rejection.

Various income and price elasticities based on the first and second stage estimates are reported in Tables 4a and 4b. Note that since $x_G = V_G Q_G$ expenditure responses of the sorts implied by the estimated Engel curves (the familiar Marshallian elasticities) combine both quantity and unit value responses. We therefore report separate quantity and unit value elasticities, but since the latter turn out to be very small, the quantity elasticities are almost the same as the usual Marshallian elasticities. Given our specification the decomposition is simple: a proportion $(1/1 + b_G)$ of price and expenditure responses is due to quantity changes and $(b_G/1 + b_G)$ to unit value changes. Note the implication that if b_G is imprecisely estimated or if b_G and β_G are highly correlated, then even if total spending significantly affects a budget share we may not be able to reject a unit budget elasticity for quantity — this is evident, for instance, in the clothing equation. A surprise is perhaps the very low budget elasticity of the vegetable and fruit category, which is at variance with results typically found for other countries.

All (uncompensated) own price elasticities are negative except that for vegetables/fruit which is insignificantly different from zero. Several significant cross price effects are observed — for instance dairy products and vegetables/fruit appear (uncompensated) complements to meat but substitutes for each other. Clothing and shoes also appear to be complementary.

The only point of reference we have in assessing these elasticities is the work of Ratinger (1995), based on monthly data on 300 household of employees from January 1990 to September 1992, apparently using published average expenditures on food. Given the differences between the two studies, the comparison is difficult. The only strong similarity is that depending on assumption and degree of aggregation of goods, Ratinger reports expenditure elasticities between .98 and 1.14 for meat.

6. Conclusion

We have presented here a new approach to the estimation of demand systems on the basis of unit values and have argued that its main advantage is consistency with demand theory[14]. Another advantage of our approach over alternatives is its relative computational simplicity, the main difference residing in the second stage where we can treat goods separately whereas, for instance, a system estimation is necessary in Deaton's approach. This simplification may allow us to consider more complicated settings, where for instance spatial patterns of consumption are of primary interest[15]. Combining, on one side, a proper treatment of the fact that quality is a choice variable and, on the other side, the spatial patterns of demand, would seem a rewarding endeavour.

[14] Monte Carlo experiments designed to compare its performance with alternative methods are presented by Lahatte et al. (1997). They do suggest that our theory-consistent specification for the log unit value equation outperforms Deaton's a first order Taylor expansion, but that both specifications perform poorly when data are generated by a more flexible form.

[15] In her work on spatial aspects to consumption, Case (1991), for instance, while aware of Deaton's work, chooses to treat unit values as error-ridden measurements of prices rather than to model them as the outcome of quality choices.

7. Appendix A: Definition of goods

7.1 Goods in equations (LHS):

The numbers refer to the documentation of the data file.

MEAT and fish: 201 pork, 202 beef, 203 other meats and innards, 204 poultry, 205 meat cans, 206 other meat products, 207 fish and fish cans

in kind: 601 pork, 603 other meat and meat products, 604 poultry

FATS and eggs, milk, cheese: 211 butter, 212 bacon and lard, 213 vegetable oil, margarine, 221 eggs, 222 milk, 233 cheese, 224 other milk products

in kind: 612 bacon and lard, 621 eggs, 622 milk

STARCHES: 231 potatoes, 241 bread, 242 bakery products, 243 wheat flour, 244 other cereal products, 245 rice, 246 pulses

in kind: potatoes

VEGETABLES and fruit: 251 fresh vegetables, 252 frozen vegetable products, 253 fresh fruit, 254 tropical fruit, 255 frozen and dried fruit

in kind: 651 fresh vegetables, 652 fresh fruit

SWEET: 261 sugar, 262 chocolate, 281 syrup and concentrates, 282 non-alcoholic beverages

ALCOHOL: 283 beer, 284 wine, 285 other alcoholic drinks

CLOTHING: 301 cloth or fabric, 302 stockings/socks, 303 knitwear for adults, 304 for children, 305 knit clothes for adults, 306 for children, 307 ready to wear clothing for men, 308 for women, 309 for children

FOOTWEAR: 313 men's, 314 women's, 315 children's shoes.

7.2 Conditioning expenditures (RHS in share equations)

The numbers refer to the documentation of the data file.

TRANSPORT and communication: 411 commuting to work, 412 other public transport, 414 telephone etc.

HYGIENE: 341 soaps, detergents, 342 cosmetics, 343 toiletries, 431 laundering, home help, dry cleaning, 432 hairdresser , cosmetics.

MEALS OUT: 291 companies canteens, 292 school canteens, 293 restaurants, 294 other catering.

CULTURE and recreation: 372 books, 373 magazines, 374 toys, 375 culture articles (durable), 376 (non durable), 377 sports equipment, 433 tourist accommodation, 434 flowers, 435 other personal services, 441 education, 442 culture, sports, entertainment, 451 recreation inside Czech Republic, 453 creche, 454 kindergarten, 455 recreation abroad, 461 other services.

ENERGY: 391 fuel all types, 402 electricity, 403 gas, 404 central heating and other municipal services.

TOBACCO: 381 tobacco products.

OTHER FOOD: 263 confectionery, 271 coffee, 272 tea, 273 ready to cook foods, 274 powdered food, 275 other food.

in kind: 675 other foods and beverages, 694 free catering.

OTHER TEXTILE: 310 textiles, 311 furs, 312 haberdashery, 316 leatherware, 421 tailor services

HOUSING: 401 rent, 422 maintenance, 531 cooperative flat payment.

in kind: 699 rent in kind

MEDICAL: 344 medicines and health care goods, 452 medical treatment.

FURNITURE and equipment: 351 furniture, 352 soft furnishings, 353 glass, porcelain, pots and pans, 354 refrigerators, freezers, 355 electric razors, hairdryers, etc., 356 washing machines, dryers.

REPAIRS: 423 cars and motorcycles repairs, 424 repairs of other goods.

8. Appendix B: Implications of the same log-linear specification for shares and log unit values

Suppose the share equations are derived from the AID functional form,

$$w_G = \alpha_G + \sum_H \delta_{GH} \ln \pi_H + \beta_G \ln X + u_G \ , \tag{8.1}$$

and that the log unit values have a similar form

$$\ln V_G = A_G + \sum_H D_{GH} \ln \pi_H + B_G \ln X + U_G \ , \tag{8.2}$$

as in Deaton (1990).

Then

$$\begin{aligned}
\partial \ln Q_G / \partial \ln \pi_H &= \partial (\ln w_G - \ln V_G) / \partial \ln \pi_H &&= \delta_{GH}/w_G - D_{GH} \ , \\
\partial \ln Q_G / \partial \ln X &= \partial (\ln w_G + \ln X - \ln V_G) / \partial \ln X &&= \beta_G / w_G + 1 - B_G \ , \\
\partial \ln \xi_G / \partial \ln \pi_H &= \partial (\ln V_G - \ln \pi_G) / \partial \ln \pi_H &&= D_{GH} - 1_{[G=H]} \ , \\
\partial \ln \xi_G / \partial \ln X &= \partial (\ln V_G - \ln \pi_G) / \partial \ln X &&= B_G \ .
\end{aligned}$$

Hence, by (2.9), considering first the case where $G \neq H$:

$$\frac{\delta_{GH} - w_G^h D_{GH}}{\beta_G + w_G^h (1 - B_G)} = \frac{D_{GH}}{B_G} \ ,$$

where both denominators are assumed different from 0. This implies

$$D_{GH} w_G^h = B_G \delta_{GH} - \beta_G D_{GH} \ .$$

For this to hold for all w_G^h and all G and $H \neq G$ requires $D_{GH} = \delta_{GH} = 0$, since $B_G \neq 0$. Thus in both equations only the own price is included. But turning to the case $G = H$, we see that the restriction is even more severe, because then for all G

$$\frac{\delta_{GG} - w_G D_{GG}}{\beta_G + w_G (1 - B_G)} = \frac{D_{GG} - 1}{B_G} \ ,$$

which implies

$$\begin{aligned}
B_G &= 1 - D_{GG} \ , \\
\beta_G &= -\delta_{GG} \ ,
\end{aligned}$$

so that in the end there is only one free slope parameter in each equation.

9. Appendix C: Details of estimation of second stage standard errors

Calling θ the vector of first stage coefficients and ϕ those from the second stage, and given the moments

$$\begin{array}{lll} G \equiv & Eg(z,\theta) & = 0 \\ H \equiv & Eh(z,\theta,\phi) & = 0 \ , \end{array}$$

$\hat{\theta}$ and $\hat{\phi}$ solve

$$\begin{array}{ll} \sum_{i=1}^{n} g(z_i,\hat{\theta}) & = 0 \\ \sum_{i=1}^{n} h(z_i,\hat{\theta},\hat{\phi}) & = 0 \ . \end{array}$$

The moment restrictions used in the first stage are

$$E\left[w_G^h - \overline{w_G}^c - \left(Z^h - \overline{Z}^c\right)\alpha_G - \beta_G\left(\ln X^h - \overline{\ln X}^c\right)\right] M^h = 0 \tag{9.1}$$

$$E\left[\ln V_G^h - \overline{\ln V_G}^c - \left(Z^h - \overline{Z}^c\right)a_G - b_G\left(\ln Q_G^h - \overline{\ln Q_G}^c\right)\right] M^h = 0 \tag{9.2}$$

$$E\left[\delta_G'\left(\zeta^c\zeta^{c\prime} - \frac{1}{n_c}\Omega_u\right) - \left(\eta_G^c\zeta^{c'} - \frac{1}{n_c}\Omega_{uv_G}\right)\right] = 0 \tag{9.3}$$

where M^h denotes either exogenous first stage regressors or predicted values of endogenous first stage regressors.

Define

$$g_{G1}^h(\alpha_G,\beta_G) \equiv \left[w_G^h - \overline{w_G}^c - (Z^h - \overline{Z}^c)\alpha_G - \beta_G\left(\ln X^h - \overline{\ln X}^c\right)\right]' M^h \tag{9.4}$$

$$g_{G2}^h(a_G,b_G) \equiv \left[\ln V_G^h - \overline{\ln V_G}^c - (Z^h - \overline{Z}^c)a_G - b_G\left(\ln Q_G^h - \overline{\ln Q_G}^c\right)\right]' M^h \tag{9.5}$$

$$g_G^h(\alpha_G,\beta_G,a_G,b_G) \equiv \begin{bmatrix} g_{G1}^{h\prime} \\ g_{G2}^{h\prime} \end{bmatrix} \tag{9.6}$$

$$h_G^h(\alpha,\beta,a,b,\delta_G) \equiv \frac{1}{n_c}\left[\delta_G'\left(\zeta^c\zeta^{c\prime}-\tfrac{1}{n_c}\hat{\Omega}_u\right)-\left(\eta^c\zeta^{c\prime}-\tfrac{1}{n_c}\hat{\Omega}_{uvG}\right)\right] \tag{9.7}$$

Then $(\hat{\alpha},\hat{\beta},\hat{c},\hat{b},\hat{\delta})$ solve

$$\sum_{h=1}^{n} g_G^h\left(\hat{\alpha}_G,\hat{\beta}_G,\hat{c}_G,\hat{b}_G\right)=0 \tag{9.8}$$

$$\sum_{h=1}^{n} h_G^h\left(\hat{\alpha},\hat{\beta},\hat{c},\hat{b},\hat{\delta}_G\right)=0 \tag{9.9}$$

for all G.

To describe asymptotic standard error formulae, define $G_\theta \equiv \partial G/\partial\theta'$, $H_\theta \equiv \partial H/\partial\theta'$, $H_\phi \equiv \partial H/\partial\phi'$ and $V = V([g',h'])$.Then

$$V_{asy}\left[\sqrt{n}\left(\hat{\theta}-\theta\right)\right] = G_\theta^{-1}V_{gg}G_\theta^{-1\prime}$$

and it is easily seen, by first-order Taylor expansion, that

$$V_{asy}\left[\sqrt{n}\left(\hat{\phi}-\phi\right)\right] = H_\phi^{-1}V_{hh}H_\phi^{-1\prime} + H_\phi^{-1}H_\theta G_\theta^{-1}V_{gg}G_\theta^{-1\prime}H_\theta' H_\phi^{-1\prime} - H_\phi^{-1}\left[H_\theta G_\theta^{-1}V_{gh}+V_{hg}G_\theta^{-1\prime}H_\theta'\right]H_\phi^{-1\prime}.$$

$$Cov_{asy}\left[\sqrt{n}\left(\hat{\phi}-\phi\right),\sqrt{n}\left(\hat{\theta}-\theta\right)\right] = G_\theta^{-1}\left[V_{gh}-V_{gg}G_\theta^{-1\prime}H_\theta'\right]H_\phi^{-1\prime}.$$

Standard error estimates are calculated using

$$\hat{V} = \begin{bmatrix} \hat{V}_{gg} & \hat{V}_{gh} \\ \hat{V}_{gh}' & \hat{V}_{hh} \end{bmatrix} = \begin{bmatrix} \sum_{h=1}^{n} g_G^h g_G^{h\prime} & \sum_{h=1}^{n} g_G^h h_G^{h\prime} \\ \sum_{h=1}^{n} h_G^h g_G^{h\prime} & \sum_{h=1}^{n} h_G^h h_G^{h\prime} \end{bmatrix}. \tag{9.10}$$

Estimates of derivatives H_θ are calculated numerically, whereas analytical computation of G_θ and H_ϕ is straightforward.

10. Appendix D: Figures

Figure 1(a,i) , Quantity and Total Expenditure: Meat

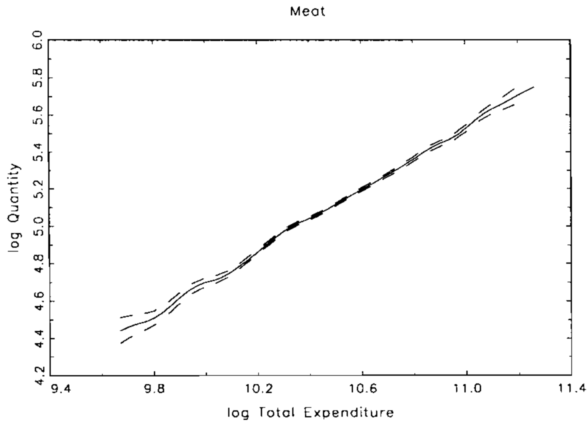


Figure 1(a,ii) , Unit Value and Total Expenditure: Meat

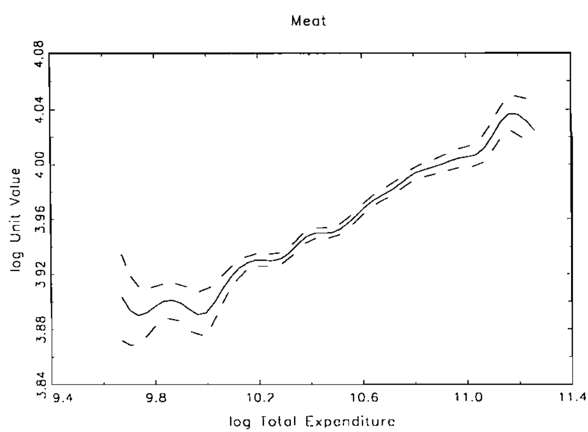


Figure 1(a,iii) , The Engel Curve: Meat

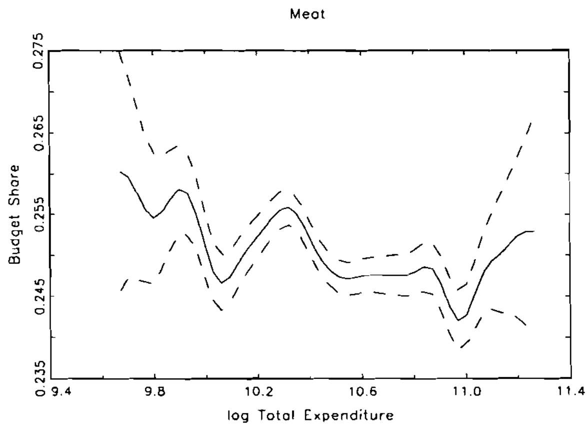


Figure 1(a,iv) , Quantity and Unit Value: Meat

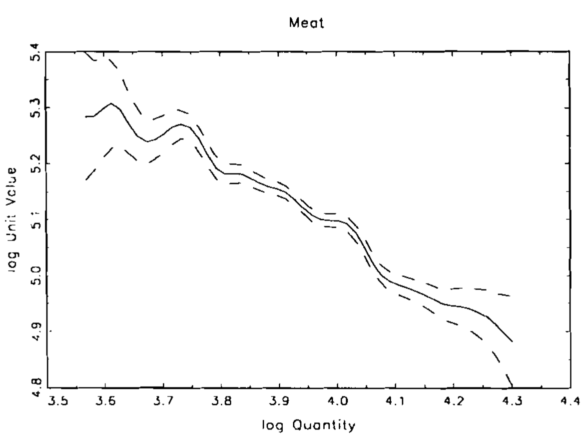


Figure 1(b,i) , Quantity and Total Expenditure: Alcohol

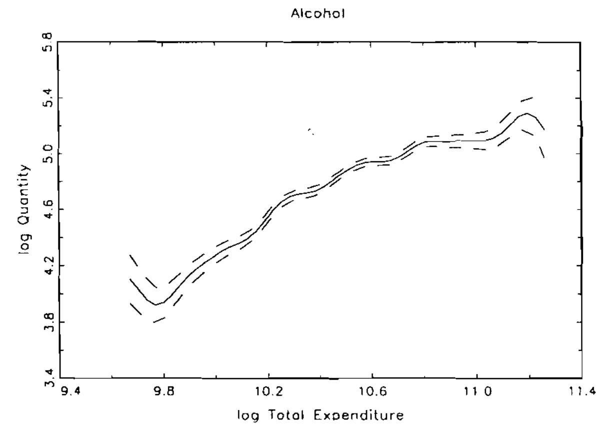


Figure 1(b,ii) , Unit Value and Total Expenditure: Alcohol

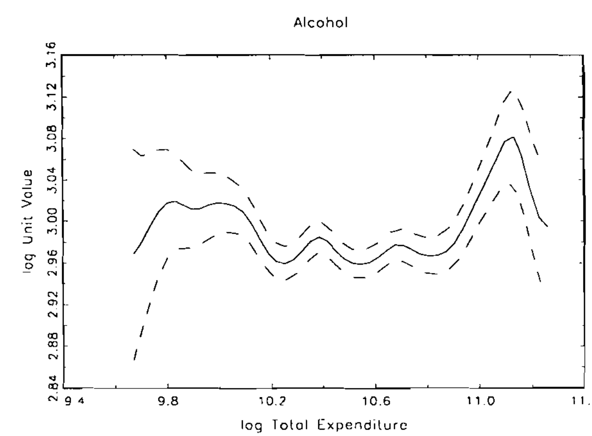


Figure 1(b,iii) , The Engel Curve: Alcohol

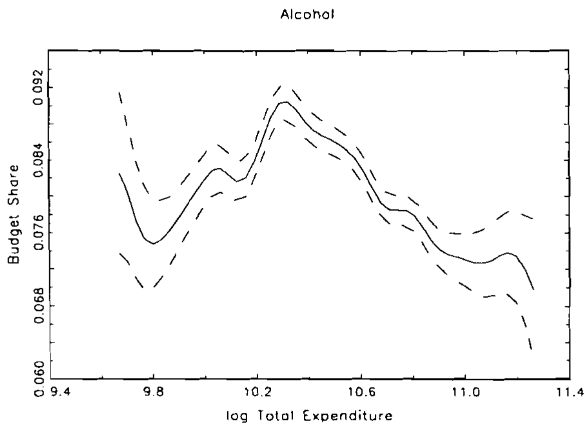


Figure 1(b,iv) , Quantity and Unit Value: Alcohol

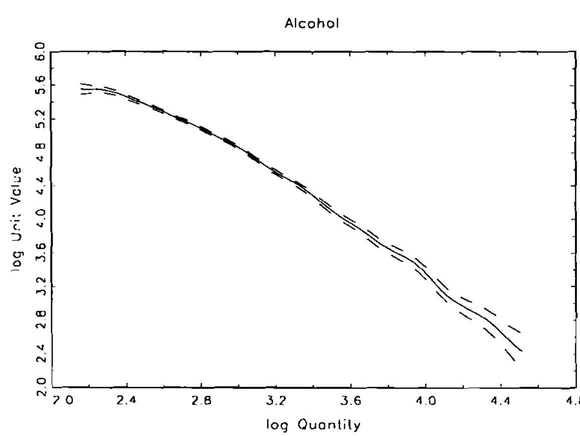


Figure 1(c,i) , Quantity and Total Expenditure: Clothing

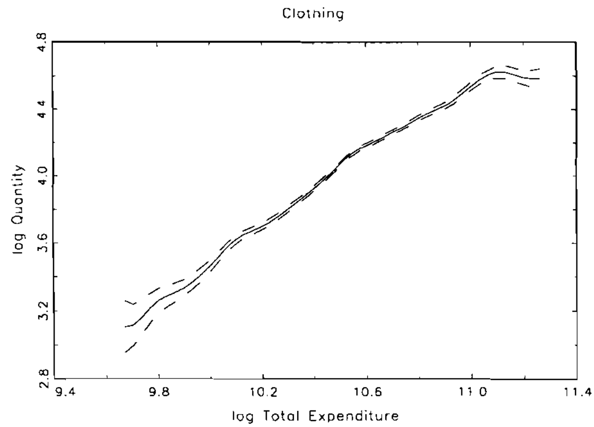


Figure 1(c,ii) , Unit Value and Total Expenditure: Clothing

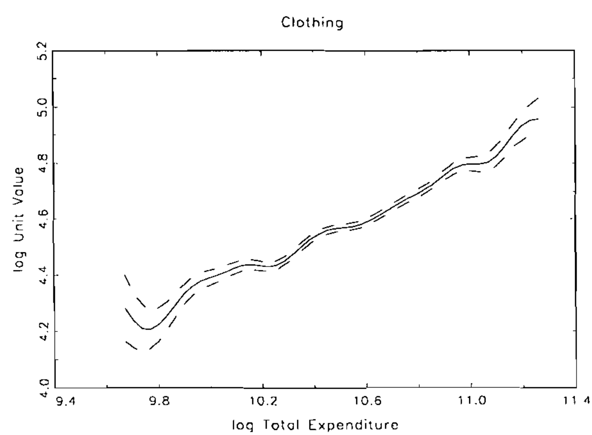


Figure 1(c,iii) , The Engel Curve: Clothing

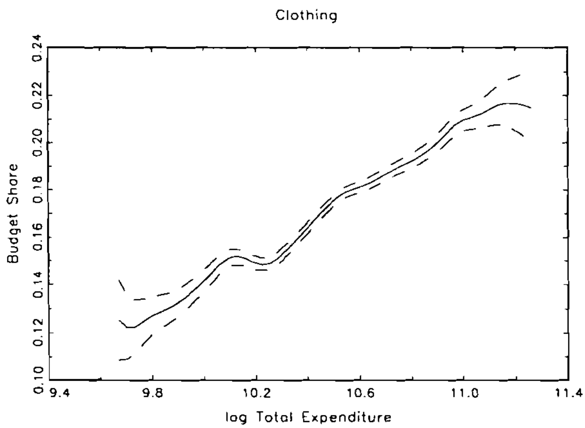
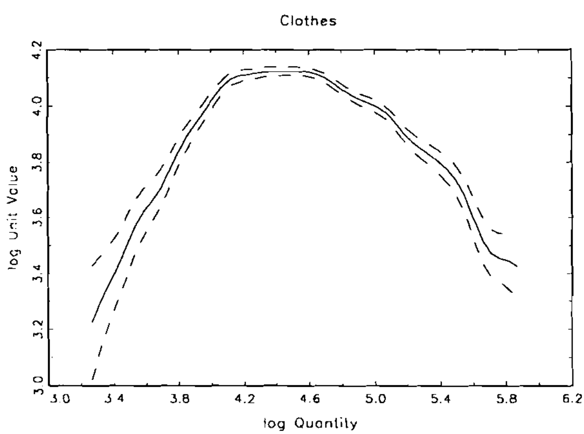


Figure 1(c,iv) , Quantity and Unit Value: Clothing



11. Appendix E: Estimates

11.1 Stage 1 parameters

Table 1a: Unit value equations: all coefficients × 100 (continued opposite)

	Meat		Dairy		Starches	
	coef.	*t*	coef.	*t*	coef.	*t*
Household characteristics						
Wife's participation	**4.654**	*4.2*	1.839	*0.7*	**6.718**	*3.8*
Blue collar	-.718	*-1.6*	**-3.005**	*-2.8*	**-1.453**	*-2.0*
Farmer	-.741	*-1.2*	**-7.256**	*-5.2*	**-2.461**	*-2.6*
Age of head of household	-2.810	*-1.6*	**-10.442**	*-2.8*	**-12.258**	*-5.3*
Age of hoh squared	.189	*1.1*	.745	*1.9*	**1.197**	*5.0*
Owner-occupier	-.318	*-.5*	.994	*.9*	**-2.298**	*-3.4*
No mod-cons	**-1.267**	*-2.1*	-.181	*-.2*	.023	*.0*
Number of hh members	-.421	*-1.2*	1.024	*1.3*	**1.789**	*3.6*
Average age of children	-.059	*-1.3*	-.101	*-1.0*	-.103	*-1.5*
Basic education - hoh	-.746	*-1.7*	**-2.318**	*-2.1*	-.460	*-.7*
Advanced education - hoh	**2.537**	*5.0*	**6.086**	*4.8*	**2.692**	*3.1*
Rural	-.222	*-.5*	-1.558	*-1.3*	**-2.351**	*-3.1*
Space per person	-.013	*-.5*	**-.215**	*-4.3*	**-.074**	*-2.3*
Durable ownership						
Gas supplied	**1.354**	*2.9*	-.353	*-.3*	.558	*.8*
Number of leisure durables	-.009	*-.1*	.216	*.9*	**.333**	*2.2*
Freezer	**-.904**	*-2.4*	-.266	*-.3*	**-2.497**	*-4.4*
Phone	-.089	*-.2*	1.195	*1.3*	.627	*1.1*
Car or motor bike	-.836	*-1.9*	**-2.888**	*-2.7*	-1.148	*-1.7*
Automatic washing machine	.500	*1.2*	**2.204**	*2.3*	.513	*.8*
Food processor	-.370	*-1.0*	-.967	*-1.2*	-.220	*-.4*
Caravan and/or dacha	**-1.058**	*-2.3*	-.877	*-.8*	.291	*.4*
Garage	.818	*1.6*	-.014	*.0*	-.722	*-1.2*
ln(Quantity)	1.471	*.9*	**-9.683**	*-2.2*	**-6.917**	*-2.3*

Table 1a: Unit value equations: all coefficients × 100 (end)

Veg/Fruit		Sweet		Alcohol		Clothes		Shoes	
coef.	*t*	coef.	*t*	coef.	*t*	coef.	*t*	coef.	*t*
692	.3	**8.233**	2.4	**23.292**	5.4	**38.080**	6.8	**46.318**	7.0
-.080	-.1	1.027	.7	-3.233	-1.7	**-4.606**	-2.0	-2.920	-1.2
-6.158	-4.0	.469	.2	**-7.787**	-2.9	-4.414	-1.1	-4.780	-1.0
-12.370	-3.0	-8.339	-1.5	**-34.287**	-4.4	**-34.316**	-2.7	-12.784	-.8
.870	2.2	.928	1.7	**3.239**	4.6	**3.500**	3.4	1.603	1.3
-2.127	-1.9	**-3.019**	-2.0	**-5.604**	-3.0	.639	.2	**-4.502**	-1.6
-1.339	-1.3	-2.271	-1.5	**-3.460**	-2.0	1.618	.7	-.000	-.0
-1.935	-1.8	**3.580**	2.4	1.656	.6	-5.448	-1.3	**-12.176**	-2.5
-.500	-5.0	-.128	-.9	-.345	-1.9	.189	.9	**1.438**	6.3
-.425	-.4	1.038	.7	**-4.956**	-2.4	-4.012	-1.3	-3.536	-1.0
2.279	1.7	-1.013	-.5	**6.444**	2.6	6.536	1.7	2.712	.6
-2.846	-2.4	-1.726	-1.1	-3.337	-1.8	-.879	-.4	-4.552	-1.8
-.112	-2.0	.070	1.0	-.006	-.1	-.142	-1.1	-.021	-.1
989	1.0	-.960	-.7	2.593	1.4	1.811	.8	1.755	.7
-.317	-1.4	.461	1.5	**.949**	2.4	.746	1.6	.046	.1
-5.513	-6.8	**-2.628**	-2.1	**-2.990**	-2.0	.680	.3	-.542	-.2
156	.2	-1.210	-1.0	1.158	.8	-2.425	-1.4	-.825	-.5
-1.725	-1.7	1.296	.9	-2.871	-1.6	**5.359**	2.7	-.014	-.0
997	1.1	.601	.5	2.513	1.5	-1.650	-.8	-1.034	-.5
-2.642	-3.5	**3.043**	2.9	**2.752**	2.2	-.700	-.5	-.451	-.3
-1.202	-1.2	.725	.5	**-6.362**	-3.6	-1.168	-.6	-.766	-.3
-2.112	-2.5	.337	.3	-2.367	-1.6	2.860	1.7	.107	.1
12.365	1.9	-10.059	-1.1	-3.219	-.2	20.326	.7	5.888	.2

Table 1b: Engel curves: all coefficients × 100 (continued opposite)

	Meat		Dairy		Starches	
	coef.	*t*	coef.	*t*	coef.	*t*
Household characteristics						
Wife's participation	**-3.980**	*-2.8*	-1.254	*-1.3*	**-2.076**	*-2.7*
Blue collar	**.819**	*3.3*	.170	*1.0*	.157	*1.2*
Farmer	**1.622**	*5.2*	.173	*.8*	**.388**	*2.3*
Age of head of household	**5.362**	*6.7*	.948	*1.7*	**1.380**	*3.5*
Age of hoh squared	**-.494**	*-5.5*	-.028	*-.5*	**-.091**	*-2.0*
Owner-occupier	**.726**	*2.9*	**.554**	*3.2*	**.338**	*2.6*
No mod-cons	-.178	*-.7*	.202	*1.1*	.169	*1.2*
Number of. hh members	**-1.070**	*-2.6*	**1.842**	*6.7*	**1.002**	*4.6*
Average age of children	.051	*1.9*	.020	*1.1*	**.063**	*4.5*
Basic education - hoh	**.831**	*3.8*	**-.569**	*-3.7*	219	*1.9*
Advanced education - hoh	**-1.433**	*-4.9*	-.172	*-.8*	**-.386**	*-2.6*
Rural	**.571**	*2.0*	-.037	*-.2*	.290	*1.9*
Space per person	**.025**	*2.1*	.013	*1.6*	.012	*1.7*
Durable ownership						
Gas supplied	.140	*.6*	**-.717**	*-4.2*	**-.299**	*-2.3*
Number of leisure durables	.039	*.7*	**-.092**	*-2.4*	003	*.1*
Freezer	**.593**	*3.0*	-.092	*-.7*	**-.261**	*-2.6*
Phone	**.759**	*3.4*	**-.363**	*2.4*	-.202	*-1.8*
Car or motor bike	.101	*.4*	-.189	*-1.1*	-.202	*-1.5*
Automatic washing machine	-.317	*-1.5*	**.313**	*2.0*	-.057	*-.5*
Food processor	-.191	*-1.0*	**.280**	*2.2*	**-.201**	*-2.2*
Caravan and/or dacha	.093	*.4*	-.202	*-1.2*	**.267**	*2.1*
Garage	.299	*1.4*	-.045	*-.3*	-.196	*-1.8*
Conditioning expenditures						
ln(Transport)	**-.693**	*-3.9*	**.339**	*2.4*	.008	*.1*
ln(Hygiene)	**-1.957**	*-4.4*	**.878**	*2.9*	**-.977**	*-4.2*
ln(Food out)	-.042	*-.2*	**-.323**	*-2.2*	.046	*.4*
ln(Culture)	**-1.201**	*-5.4*	**-.344**	*-2.2*	-.065	*-.6*
ln(Fuel)	.149	*.8*	.261	*1.7*	**.287**	*2.2*
ln(Tobaccco)	**.246**	*4.6*	**-.223**	*-5.9*	**-.077**	*-2.9*
ln(Other food)	**-.636**	*-2.0*	**.543**	*2.4*	-.054	*-.3*
ln(Textiles)	**-.900**	*-7.6*	**-.275**	*-3.4*	**-.260**	*-4.3*
ln(Medical)	**-.461**	*-5.9*	**.128**	*2.5*	**.158**	*3.9*
ln(Furniture)	**-.352**	*-4.7*	-.087	*-1.6*	-.075	*-1.9*
No food out	1.104	*.7*	**-2.392**	*-2.2*	1.031	*1.0*
No Tobacco	.461	*1.0*	**-.735**	*-2.3*	**-.773**	*-3.3*
No Medical	-.905	*-1.2*	-.040	*-.1*	.472	*1.3*
ln(Total expenditure)	**5.305**	*2.3*	**-8.552**	*-5.5*	-2.276	*-1.9*

Table 1b: Engel curves: all coefficients × 100 (end)

Veg/Fruit		Sweet		Alcohol		Clothes		Shoes	
coef.	*t*	coef.	*t*	coef.	*t*	coef.	*t*	coef.	*t*
1.234	1.6	.081	.1	-1.614	-1.3	**5.914**	3.8	**1.695**	2.5
-.095	-.8	-.022	-.3	-.261	-1.2	-.495	-1.7	**-.273**	-2.3
-.305	-2.1	-.070	-.6	.249	1.1	**-1.541**	-4.3	**-.516**	-3.6
055	.1	.237	.9	**1.835**	2.6	**-7.872**	-9.0	**-1.945**	-5.2
060	1.3	.007	.2	**-.255**	-3.3	**.638**	6.7	**.163**	4.0
-.519	-4.3	-.159	-1.8	-.086	-.4	**-.872**	-3.1	.018	.2
.274	2.0	-.039	-.4	-.149	-.7	-.359	-1.2	.081	.6
.422	2.0	.229	1.5	**-1.631**	-4.7	**-1.299**	-2.8	**.506**	2.7
-.026	-1.9	-.019	-1.9	**-.090**	-4.1	.002	.1	-.001	-.1
060	.5	.099	1.3	.022	.1	**-.682**	-2.7	.019	.2
-.012	-.1	-.098	-.9	.034	.1	**1.750**	5.0	**.316**	2.3
-.486	-3.6	.077	.8	**.571**	2.6	**-.800**	-2.6	-.187	-1.4
-.013	-2.2	**.011**	2.2	-.005	-.5	**-.035**	-2.5	-.008	-1.6
-.162	-1.3	**-.222**	-2.4	.267	1.2	**.680**	2.4	**.312**	2.5
-.017	-.6	.002	.1	**.152**	3.3	-.027	-.4	**-.060**	-2.3
043	.4	-.076	-1.1	.277	1.7	-.361	-1.6	-.124	-1.3
-.041	-.3	-.029	-.4	.013	.1	-.183	-.7	.045	.4
-.109	-.8	.052	.6	-.308	-1.4	**.662**	2.3	-.006	.0
006	.1	-.003	-.0	-.010	-.1	-.022	-.1	.089	.8
099	1.1	.001	.0	**-.494**	3.3	.372	1.8	.133	1.5
-.281	-2.2	-.156	-1.7	**.481**	2.2	-.167	-.6	-.035	-.3
-.009	-.1	-.073	-1.0	**-.422**	-2.4	.348	1.4	.097	1.0
193	1.9	.034	.5	-.209	-1.3	.323	1.5	006	.1
1.063	4.4	**.536**	3.4	**-1.016**	-2.7	.986	1.9	**.487**	2.4
-.128	-1.2	-.018	-.2	**.491**	2.8	-.094	-.4	.068	.7
-.003	-.0	.059	.8	**.419**	2.4	**.938**	3.8	**.197**	2.0
-.021	-.2	.104	1.7	.036	.2	**-.697**	-2.8	-.118	-1.0
-.150	-5.4	-.023	-1.2	**.293**	6.5	-.064	-1.0	-.003	-.1
1.087	6.2	**.363**	3.0	-.120	-.4	**-1.053**	-2.7	-.129	-.8
094	1.6	**-.132**	-3.2	**-.250**	-2.4	**1.466**	10.2	**.258**	4.2
.222	5.8	**.108**	4.0	-.088	-1.4	-.028	-.3	-.038	-1.0
-.030	-.7	-.040	-1.5	.067	1.0	**.437**	4.8	**.080**	2.0
-1.116	-1.4	.405	.7	2.147	1.7	-1.083	-.6	-.097	-.1
-.543	-2.1	.065	.4	.671	1.8	.793	1.6	.061	.3
047	.1	.483	1.8	-.050	-.1	.260	.3	-.268	-.8
-4.846	-4.1	-1.364	-1.6	**4.294**	2.2	**8.017**	3.1	-.578	-.5

11.2 Stage 2 parameters

Table 2: Unrestricted Estimates of δ: all entries $\times$ 100

	Meat	Dairy	Starches	Veg/Fruit	Sweet	Alcohol	Clothes	Shoes
Meat	2.143	**5.174**	**16.676**	-2.571	**5.264**	**4.213**	**-28.681**	-2.218
	6.128	*.829*	*2.042*	*1.953*	*.530*	*660*	*1.361*	*2.169*
Dairy	-2.183	**5.508**	-1.361	2.778	-.053	**-2.000**	**-3.128**	.439
	4.947	*.885*	*1.727*	*1.546*	*.365*	*253*	*.608*	*.929*
Starches	-5.438	**3.335**	**7.235**	-.241	**2.698**	**-4.480**	**-2.928**	-.182
	4.278	*.794*	*1.853*	*1.402*	*.417*	*275*	*.835*	*1.367*
Veg/Fruit	.596	**-1.963**	**-7.247**	**5.242**	**-3.101**	**3.260**	**3.663**	-.450
	3.146	*.729*	*1.054*	*1.506*	*.335*	*297*	*.389*	*.650*
Sweet	-.147	**-2.595**	-1.407	.431	**-.976**	.030	**3.664**	**1.000**
	2.075	*.336*	*.746*	*.662*	*.455*	*.150*	*.289*	*.433*
Alcohol	.0746	-.3546	-1.179	1.742	.414	-.186	-.458	-.053
	5.713	*.5735*	*1.837*	*1.827*	*.812*	*1.781*	*1.186*	*.720*
Clothes	-.306	-2.310	-2.159	.229	-1.446	-2.365	**8.787**	-.431
	8.097	*1.463*	*3.927*	*2.731*	*1.048*	*1.526*	*3.759*	*2.034*
Shoes	.190	1.380	1.487	-.747	.016	.558	**-4.624**	1.740
	4.020	*.742*	*1.617*	*1.205*	*.279*	*324*	*.263*	*2.080*

11.3 Stage 3 parameters

Table 3: Symmetry-Restricted Estimates of γ: all entries $\times$ 100

	Meat	Dairy	Starches	Veg/Fruit	Sweet	Alcohol	Clothes	Shoes
Meat	**15.590**							
	1.100							
Dairy	**-7.321**	**8.789**						
	.466	*.412*						
Starches	.188	**-2.087**	**4.106**					
	.611	*.261*	*.550*					
Veg/Fruit	**-8.174**	**1.138**	**-2.120**	**7.860**				
	.578	*.243*	*.293*	*.641*				
Sweet	**.975**	**-1.320**	**.450**	**-.377**	-.460			
	.364	*.150*	*.190*	*.177*	*.328*			
Alcohol	.058	**.593**	**-.769**	**2.632**	**1.043**	**-2.138**		
	.301	*.139*	*.156*	*.148*	*.095*	*.509*		
Clothes	**-1.450**	.276	.238	**.601**	-.083	**-1.240**	**2.528**	
	.544	*.284*	*.363*	*.247*	*.166*	*.413*	*1.181*	
Shoes	.133	-.068	-.005	**-1.560**	**-.227**	-.178	**-.869**	**2.775**
	.257	*.116*	*.166*	*.154*	*.084*	*.098*	*.136*	*.280*

Wald test of symmetry restrictions $\chi^2_{28} = 563.96$

11.4 Elasticities

Table 4a: Quantity elasticities

$$e^Q_{GH} = \frac{\partial \ln Q_G}{\partial \ln \pi_H} = \frac{1}{1+b_G}\left(\frac{\delta_{GH}}{w_G} - 1_{[G=H]}\right); \qquad e^Q_G = \frac{\partial \ln Q_G}{\partial \ln x} = \frac{1}{1+b_G}\left(\frac{\beta_G}{w_G} + 1\right)$$

				Price					*Total*	*Mean*
	Meat	Dair.	Starc.	Veg.	Swe.	Alc.	Clot.	Sho.	*budg.*	*share*
Meat	**-.423**	**-.327**	-.017	**-.339**	.026	-.015	**-.093**	-.006	**1.195**	.250
	.018	*.031*	*.024*	*.028*	*.017*	*.018*	*.034*	*.011*	*.093*	
Dairy	**-.310**	**-.487**	**-.065**	**.109**	**-.048**	**.078**	**.104**	**.023**	**.596**	.185
	.030	*.037*	*.018*	*.019*	*.011*	*.015*	*.029*	*.008*	*.097*	
Starches	.070	**-.154**	**-.670**	**-.179**	**.054**	**-.054**	.058	.011	**.864**	.116
	.057	*.035*	*.053*	*.032*	*.019*	*.020*	*.045*	*.016*	*.115*	
Veg/Fruit	**-.767**	**.224**	**-.172**	.019	-.009	**.334**	**.158**	**-.143**	**.356**	.081
	.058	*.046*	*.034*	*.074*	*.023*	*.031*	*.046*	*.019*	*.132*	
Sweet	**.243**	**-.197**	**.112**	-.049	**-1.182**	**214**	.028	-.029	**.860**	.060
	.061	*.048*	*.039*	*.040*	*.136*	*.033*	*.050*	*.017*	*.178*	
Alcohol	-.127	-.025	**-.158**	**.285**	**.098**	**-1.345**	**-.247**	**-.051**	**1.570**	.083
	.079	*.045*	*.043*	*.052*	*.024*	*.229*	*.091*	*.022*	*.355*	
Clothes	**-.167**	**-.058**	-.033	-.002	-.027	**-.092**	**-.775**	**-.063**	**1.219**	.172
	.054	*.027*	*.020*	*.016*	*.012*	*.031*	*.211*	*.019*	*.337*	
Shoes	.049	.007	.011	**-.268**	-.034	-.023	**-.136**	**-.447**	**.842**	.053
	.033	*.043*	*.039*	*.097*	*.023*	*.028*	*.067*	*.165*	*.349*	

Table 4b: Quality elasticities: all entries × 100

$$e^{\xi}_{GH} = \frac{\partial \ln \xi_G}{\partial \ln \pi_H} = b_G e^{Q}_{GH}; \qquad e^{\xi}_{G} = \frac{\partial \ln \xi_G}{\partial \ln x} = b_G e^{Q}_{G}$$

				Price					*Total*
	Meat	Dair.	Starc.	Veg.	Swe.	Alc.	Clot.	Sho.	*budg.*
Meat	-.622	-.482	-.025	-.499	.038	-.022	-.137	-.009	1.758
	.682	*.530*	*.045*	*.548*	*.048*	*.036*	*.158*	*.019*	*1.929*
Dairy	3.000	4.717	.634	-1.059	.467	-.753	-1.011	-.224	-5.771
	1.526	*2.384*	*.362*	*.561*	*.255*	*404*	*.577*	*.136*	*3.031*
Starches	-.485	1.067	**4.633**	1.240	-.376	.372	-.403	-.074	**-5.974**
	.451	*.548*	*2.168*	*.614*	*.217*	*.221*	*.360*	*.116*	*2.869*
Veg/Fruit	**-9.480**	**2.771**	**-2.121**	.231	-.116	**4.129**	1.953	**-1.773**	4.406
	4.396	*1.392*	*1.056*	*.924*	*.293*	*1.927*	*1.059*	*.844*	*2.595*
Sweet	-2.447	1.985	-1.131	.495	11.887	-2.149	-.281	.287	-8.647
	2.549	*2.063*	*1.207*	*.642*	*12.092*	*2.198*	*.579*	*.335*	*8.921*
Alcohol	.408	.081	.510	-.919	-.315	4.329	.796	.164	-5.053
	2.103	*.440*	*2.610*	*4.701*	*1.615*	*22.151*	*4.079*	*.840*	*25.868*
Clothes	-3.394	-1.188	-.681	-.046	-.556	-1.871	-15.763	-1.274	24.773
	4.331	*1.568*	*.935*	*.333*	*.730*	*2.394*	*19.925*	*1.618*	*31.342*
Shoes	.290	.041	.064	-1.580	-.200	-.136	-.804	-2.632	4.957
	1.717	*.350*	*.444*	*9.320*	*1.188*	*818*	*4.748*	*15.529*	*29.261*

References

[1] Ayadi, M., R. Baccouche, M. Goaied and M.S. Matoussi, 1995. Spatial Variations of Prices and Household Demand Analysis in Tunisia, *paper presented at the 7th World Congress of the Econometric Society, Tokyo.*

[2] Banks, J., R.Blundell and A.Lewbel, 1996. Tax Reform and Welfare Measurement: Do We Need Demand System Estimation?, *Economic Journal*, 106, 1227-1241.

[3] Bewley, R., 1986. *Allocation Models: Specification, Estimation and Applications.* Cambridge, Mass.: Ballinger.

[4] Browning, M.J. and C.D.H. Meghir, 1991. The Effects of Male and Female Labour Supply on Commodity Demands, *Econometrica*, 59, 925-952.

[5] Case, A., 1991. Spatial Patterns in Household Demands, *Econometrica*, 59(4), 953-65.

[6] Deaton, A.S., 1987. Estimation of Own and Cross-Price Elasticities from Household Survey Data, *Journal of Econometrics*, 36, 7-3.

[7] Deaton, A.S., 1988. Quality, Quantity and Spatial Variation of Price, *American Economic Review*, 78, 418-43.

[8] Deaton, A.S., 1990. Price Elasticities from Survey Data: Extension and Indonesian Results, *Journal of Econometrics*, 44, 281-309.

[9] Deaton, A.S., 1995. *The Analysis of Household Budgets.* Oxford: Oxford University Press.

[10] Deaton, A.S. and F. Grimard, 1991. Demand Analysis for Tax Reform in Pakistan, *Discussion Paper No 151*, Research Programme in Development Studies, Centre for International Studies, Woodrow Wilson School, Princeton.

[11] Deaton, A.S. and J. Muellbauer, 1980. *Economics and Consumer Behavior.* Cambridge: Cambridge University Press.

[12] De Vreyer, P., 1996. Hétérogénéité, agrégation et demande de travail, *mimeo.*

[13] Kodde, D.A., F.C. Palm and G.A. Pfann, 1990. Asymptotic Least-Squares Estimation Efficiency Considerations and Applications, *Journal of Applied Econometrics*, 5, 229-243.

[14] Jorgenson, D.W., L.J. Lau and T.M. Stoker, 1982. The Transcendental Logarithmic Model of Aggregate Consumer Behavior, in Basman, R. and G. Rhodes (eds.) *Advances in Econometrics*, Vol.1, 97-238. Greenwich: JAI Press.

[15] Lahatte, A., F. Laisney, R. Miquel and I. Preston, 1997. Demand Systems with Unit Values: a Comparison of Two Specifications, *ZEW Discussion Paper*.

[16] Lewbel, A., 1996. Aggregation Without Separability: A Generalized Composite Commodity Theorem, *American Economic Review*, 86, 524-543.

[17] Newey, W.K., 1984. A Method of Moments Interpretation of Sequential Estimators, *Economics Letters*, 14, 201-206.

[18] Nichèle, V. and J.M. Robin, 1995. Simulation of Indirect Tax Reforms Using Pooled Micro and Macro French Data, *Journal of Public Economics*, 56(2), 225-44.

[19] Pitt, M.M., 1983. Food Preferences and Nutrition in Rural Bangladesh, *Review of Economics and Statistics*, 65, 105-14.

[20] Ratinger, T, 1995. The Own, Cross and Expenditure Elasticities for the Selected Food Groups in the Czech Republic under Economic Transition 1990-1992, *Central European Journal for Operations Research and Economics*, 1, 7-21.

[21] Stone, J.R.N., 1954. *The Measurement of Consumers' Expenditures and Behaviour in the United Kingdom*, Cambridge: Cambridge University Press.

[22] Strauss, J., 1982. Determinants of Food Consumption in Rural Sierra Leone, *Journal of Development Economics*, 11, 327-53.