Journal of Econometrics 44 (1990) 281–309. North-Holland

PRICE ELASTICITIES FROM SURVEY DATA
Extensions and Indonesian Results

Angus DEATON*

Princeton University, Princeton, NJ 08544, USA

Received June 1988, final version received May 1989

This paper extends and improves the author's earlier work on measuring own- and cross-price elasticities from spatial variation in prices using household survey data. Double-logarithmic demand functions are replaced by functions that relate budget shares to the logarithms of prices and incomes, and zero expenditures are treated appropriately. Formulae are developed for estimation and for the calculation of standard errors. Limited Monte Carlo evidence suggests that the asymptotic approximations work well in practice. An eleven-commodity system of food demands is estimated using Indonesian data from 1981.

1. Introduction

For many questions of public policy, it is important to know how consumers change their expenditures on goods in response to changes in prices. For developing countries, there are typically rather few time-series data from which price elasticities can be inferred. By contrast, cross-sectional household expenditure surveys are available for many LDC's. In Deaton (1986, 1987) I developed a methodology for using such household survey data to detect spatial variation in prices and to estimate price elasticities by comparing spatial price variation to spatial demand patterns. In the first paper I showed how to estimate the own-price elasticity for a single good by comparing its demand to its price. In the second paper the methodology was extended to cover systems of demand functions, so that cross-price elasticities could be estimated and substitution patterns studied.

Both papers, although giving satisfactory results for data from the Cote d'Ivoire, contain a number of unresolved problems. Perhaps the most serious of these is the use of double-logarithmic demand equations. Not only are such demand functions inconsistent with basic theory, but more importantly, they

*The author would like to thank Dwayne Benjamin who provided outstanding research assistance, and two referees and an associate editor for comments on a previous version. The research reported here was carried out for the Welfare and Human Resources Division of the World Bank. The World Bank does not accept responsibility for the views expressed herein, which are those of the author and should not be attributed to the World Bank or to its affiliated organizations.

0304-4076/90/$3.50

cannot be used to model households that do not purchase all goods. As a result, many sample points have to be deleted prior to estimation. Even if the selection did not introduce bias, the estimated demand functions are thereby conditioned on positive consumption. However, for most policy purposes, it is the *unconditional* demands that are of interest. The revenue effects of a tax change depend on how total demand is altered and not on whether changes take place at the extensive or intensive margins.

In the published version of the first paper, Deaton (1988), the logarithmic formulation was abandoned, but the analysis was confined to the single-commodity case, and thus to measurement of own-price responses. The current paper provides a unified statement of the methodology for the system case without using double-logarithmic forms and is intended to supercede previous treatments. The first part of the paper summarizes the model and describes the estimation procedure. A brief Monte Carlo experiment illustrates how the procedure works in practice as well as the consequences of following alternative estimation strategies of the kind that have appeared in the literature. The second part of the paper presents an application of the new version of the model to data from the 1981 SUSENAS Indonesian household survey. Results are given for a moderately large demand system of eleven commodities. Another application of the same techniques can be found in Laraki's (1988) analysis of food subsidies and demand patterns in Morocco. The appendix derives standard errors for the estimators presented in section 2. A SAS computer program that implements these procedures is available from the author or from the Welfare and Human Resources Division of the World Bank.

2. Model formulation and estimation

The framework for the analysis is a model of consumer behavior in which households choose how much of a commodity to buy and in what quality or grade. Commodities are considered as collections of heterogeneous goods within which consumers can choose more or less expensive items, so that the unit value of a commodity, the price paid per physical unit, is a matter of choice. Both quantity and quality choice are functions of household income, household characteristics, and price. Prices of any one good will typically affect the quantities and qualities chosen of all goods.

To estimate such a model, data are required on household expenditures on a range of goods as well as on physical quantities purchased. Weight may well be the most natural measure of physical quantity, but it is by no means the only one. For example, it might be convenient to consider a commodity 'flour and flour products' that contained purchases of both flour and bread. In such a case, it would not be sensible to add kilos of bread to kilos of flour, but rather to convert bread to its flour equivalent before adding. Calories may also be a convenient common unit in which to aggregate. While all household budget surveys collect data on household expenditures, not all ask questions about physical quantities. The methods of this paper cannot be applied to survey data without such information.

The second major data requirement, and one that is satisfied by virtually all household surveys, is that households be geographically 'clustered' within the sample. Survey organizations nearly always adopt such a design because it minimizes transport costs and allows a group of households to be interviewed at the same time. For current purposes, clustering is important because it means that households within each cluster can be assumed to face the same prices for market goods. Note that the validity of the 'same price' assumption requires not only the geographical proximity of the households, but also that they be interviewed at approximately the same time.

The model that I shall work with is as follows. For household i in cluster c there are two equations for good G:

$$w_{Gic} = \alpha_G^0 + \beta_G^0 \ln x_{ic} + \gamma_G^0 \cdot z_{ic} + \sum_{H=1}^{N} \theta_{GH} \ln p_{Hc} + \left(f_{Gc} + u_{Gic}^0 \right), \qquad (1)$$

$$\ln v_{Gic} = \alpha_G^1 + \beta_G^1 \ln x_{ic} + \gamma_G^1 \cdot z_{ic} + \sum_{H=1}^{N} \psi_{GH} \ln p_{Hc} + u_{Gic}^1. \qquad (2)$$

In eq. (1), w_{Gic} is the budget share of good G in household i's budget, defined as expenditure on the good divided by total expenditure on all goods and services, x_{ic}. This share is assumed to be a linear function of the logarithm of total expenditure, of the logarithms of the prices of all of the N goods, and of a vector of household characteristics z_{ic}. The remaining terms are f_{Gc}, a cluster-fixed effect for good G, and an idiosyncratic residual u_{Gic}^0. The second equation relates to the unit value of good G, v_G, as defined as the expenditure on the good divided by the quantity bought. The logarithm of the unit value is a function of the same variables that appear in the share equation, with the exception of the cluster-fixed effect. The basic idea is that the logarithm of unit value is the logarithm of quality plus the logarithm of price. In consequence, if there were no quality effects, unit values would move proportionally with price. However, prices, income, and characteristics all affect the choice of quality and so all appear in the unit-value equation. Although it is difficult to rule out the possibility of cluster-fixed effects in the unit-value equation, the model depends on their exclusion; since prices are not measured, the identification of the model requires a direct link between prices and unit values, a link that would be broken by the presence of fixed effects.

One nonstandard feature of the eqs. (1) and (2) is that the prices for the goods, $\ln p_{Gc}$, are not observed, so that it is not possible to estimate the equations directly. Note that prices are assumed to be the same for all households in cluster c, so that there is no i suffix on these variables. The budget share in eq. (1) is observed for all households, but the unit value for good G in eq. (2) is observed only for those households that record at least one purchase in the market for that commodity. Households with zero budget shares do not generate a corresponding unit value, just as in the labor supply literature, individuals who do not work do not have recorded wage rates. However, in the current case there will also be households with a positive budget share but for whom there is no recorded unit value. This occurs if households consume own-produced goods whose value has to be imputed. Different surveys will generally have different imputation rules, and it will only rarely be the case that the prices used for imputation can be treated as genuine observations on unit values. For each cluster c, I shall denote by n_c the number of households in the cluster, and by n_{cG}^{+} the number of households that have observations on both the budget share and the unit value of good G.

Eq. (1) looks very much like the 'Almost Ideal' demand system of Deaton and Muellbauer (1980) in which budget shares are a linear function of the logarithms of real expenditure and prices. However, there are a number of reasons why the model here is different. First, eqs. (1) and (2) should not be regarded as a direct representation of preferences, but simply as the regression functions of budget shares and unit values conditional on the included right-hand-side variables. Zero expenditures are included, so that the conditional expectation is taken over purchasers and nonpurchasers alike. There is no guarantee that there exist preferences that generate a regression function like (1). Instead of following the traditional methodology of postulating a linear structural model and then dealing with the zero censoring separately and subsequently, I am directly postulating that the conditional expectations or regression functions take the form (1) and (2). Such a procedure has several advantages. As argued in the introduction, it is this regression function that is of interest for policy, and estimation of the underlying structural model is not required. Estimation is simplified because I am dealing with a linear model and not with a Tobit or its multivariate generalization. Further, it is the regression function which is identified from the data, and it is far from clear whether it is possible to disentangle the effects of the censoring from the underlying functional form without essentially arbitrary and untestable identifying assumptions. Of course, there still remains unresolved the question of whether (1) is a plausible functional form for the budget shares, given that the zero observations are included.

The second reason why (1) and (2) differ from a model like the almost ideal demand system is that we are no longer within the framework of a standard demand model where quantities are a function of prices and the budget. Here, consumers choose both quantity and quality, so that expenditure is the product not only of quantity and price, but of quantity, quality, and price. As a consequence, the analysis must take account of price and income elasticities of quality, and the existence of these effects also complicates the relationship between the parameters and the elasticities that we are ultimately interested in measuring.

The parameters β_G^0 in (1) and β_G^1 in (2) determine the total expenditure elasticities of quantity and quality. Since $\beta_G^1 = \partial \ln v_G / \partial \ln x$ and since unit value is price multiplied by quality, the parameter is simply the expenditure elasticity of quality. If (1) is differentiated with respect to $\ln x$ and ε_G is the (quantity) demand elasticity, we have

$$\partial \ln w_G / \partial \ln x = \beta_G^0 / w_G = \varepsilon_G + \beta_G^1 - 1, \tag{3}$$

since the logarithm of the share is the sum of the logarithms of quantity and quality less the logarithm of expenditure. Rearranging,

$$\varepsilon_G = \left(1 - \beta_G^1\right) + \left(\beta_G^0 / w_G\right). \tag{4}$$

Turning to the price elasticities, ψ_{GH} is the matrix of own- and cross-price elasticities of the unit values; if price were to have no effect on quality, Ψ would be the identity matrix. In general, the elasticities of quality with respect to price are $\psi_{GH} - \delta_{GH}$, for Kronecker delta δ_{GH}. If ε_{GH} is the standard matrix of own- and cross-price elasticities of quantities, then, differentiating (1) with respect to $\ln p_H$, we have

$$\partial \ln w_G / \partial \ln p_H = \varepsilon_{GH} + \psi_{GH} = \theta_{GH} / w_G, \tag{5}$$

so that

$$\varepsilon_{GH} = -\psi_{GH} + \theta_{GH} / w_G. \tag{6}$$

It is the estimation of the quantities ε_{GH} and ε_G to which I give the most attention in what follows, although note that, for some purposes, interest might focus on the income and price elasticities of quantity and quality together, i.e., on $\varepsilon_G + \beta_G^1$ and $\varepsilon_{GH} + \psi_{GH} - \delta_{GH}$, quantities that are easily calculated if required. Note that the elasticity formulae typically contain both parameters and data and will therefore vary across the sample. I shall typically ignore this variation and evaluate formulae at the sample means of the data.

Given that prices are not observed, all of the parameters cannot be estimated without further prior information. The basic result that yields identification is a formula that links the effects of prices on quality choice to conventional price and total expenditure elasticities. Given a separability assumption about the basic goods that comprise each heterogeneous commodity, it is shown in Deaton (1988) that

$$\psi_{GH} = \delta_{GH} + \beta_G^1 \varepsilon_{GH} / \varepsilon_G. \tag{7}$$

According to (7), the price of good H only effects the quality of good G to the extent that there is a cross-price quantity elasticity ε_{GH}. Given such an elasticity, the effect works through the change in the total quantity of good G; β_G^1/ε_G is the elasticity of quality of G with respect to total expenditure on G.

Assuming that (7) holds at the sample means, (4) and (6) can be used to substitute for ε_{GH} and ε_G in (7), and we obtain a relationship linking the underlying parameters:

$$\psi_{GH} = \delta_{GH} + \frac{\beta_G^1(\theta_{GH}/w_G - \psi_{GH})}{(1-\beta_G^1) + \beta_G^0/w_G}. \tag{8}$$

It is convenient to define the vector ξ by

$$\xi_G = \beta_G^1 / \left\{ (1-\beta_G^1) w_G + \beta_G^0 \right\}, \tag{9}$$

so that, in matrix notation, (8) becomes

$$\Psi = I + D(\xi)\Theta - D(\xi)D(w)\Psi, \tag{10}$$

where I is the $(N \times N)$ identity matrix and $D(x)$ denotes a diagonal matrix with the vector x on its diagonal.

I am now in a position to discuss the method of estimation. There are both important analogies and important differences between the methods used here and the methods of estimation routinely used for panel data. In panel data, we typically have a short time series on a large cross-section of individuals. Error structures are specified that allow either fixed or random effects for each individual, and estimators are sought that will be consistent as the number of individuals in the cross-section increases with the number of time-series observations held constant. In the current application, the role of the individuals is taken by the clusters or villages in the survey, and the repeat, time-series observations are replaced by the individual households within each cluster. In sample surveys, the cluster size does not vary very much across surveys, sample sizes, or countries, and is usually between six and fifteen households. In consequence, as in the case of panel data, I require estimators that are consistent as the number of clusters increases, holding cluster size constant. As far as error specification is concerned, eq. (1) assumes that there is a fixed effect in the demand equation. This allows each household in each cluster to share a idiosyncratic cluster effect that could represent shared preferences (villages are often homogeneous with respect to race, tribe, religion, or occupation), or weather, or distance, or many other factors. Such a specification allows for arbitrary patterns of spatial correlations in consumption patterns; the spatial autocorrelation models estimated by Case (1988), and shown by her to be important, are special cases of the model estimated here.

It is very important that the fixed effects be permitted to be correlated with the included exogeneous variables, particularly income, and it is this feature that rules out the use of a 'random' effect for each cluster and the associated error components model. For the unit-value equation (2), the addition of a fixed effect would destroy identification, because the unit values would no longer give any useful information about the prices. However both random errors u^0 and u^1 in (1) and (2) are allowed to have cluster components; in particular, there is no supposition that measurement error cancels out over the households in each cluster. When there are only a few sample households in the village, the average measurement error for the cluster will typically have lower variance than the measurement error for each household, but that does not mean that it can be ignored. I note finally that while in most panel data sets there are the same number of time-series observations on each individual, it is typically not the case here that there are the same number of sample households per cluster; even when intended by design, nonresponse always introduces some variation. This variation introduces some additional complexity into the algebra but raises no issues of principle.

The estimation takes place in two stages. At the first, eqs. (1) and (2) are estimated equation by equation by OLS with cluster means subtracted from all data. The subtraction of cluster means removes not only the fixed effects in (1) but also the cluster invariant prices in both equations. The resulting 'within' estimates of β_G^0, γ_G^0, β_G^1, and γ_G^1 are consistent in spite of the lack of information on prices and fixed effects. Denote these parameter estimates as $\tilde{\beta}_G^0$, $\tilde{\gamma}_G^0$, $\tilde{\beta}_G^1$, and $\tilde{\gamma}_G^1$. Although I shall not be dealing here with a 'complete set' of demand equations, in which the budget shares of the goods add to unity, these parameter estimates respect the adding-up conditions in the sense that for a complete system the vectors of parameter estimates $\tilde{\beta}^0$ and $\tilde{\gamma}^0$ add to zero. This result is a consequence of the fact that for a system of OLS estimators $B = (X'X)^{-1}(X'Y)$, where Y is the matrix of observations on the budget shares, $Y\iota = X\gamma$ implies $B\iota = \gamma$ for vector of units ι. In this case the budget shares add to unity, so that γ is zero except for the element corresponding to the constant term in X. All this only works if zero observations on consumption are *included*. Otherwise, the sample of included households is different for each commodity, we no longer have a simple multivariate regression set, and the result does not apply.

Note that, from (4) and (9), the estimates of the total expenditure elasticities of quantity and quality as well as the parameters ξ are functions only of these first-stage parameters. Denote the residuals from the two sets of regressions as e_{Gic}^0 and e_{Gic}^1. These can be used to give consistent estimates of the variances and covariances of the residuals in (1) and (2) as follows:

$$\tilde{\sigma}_{GH}^{00} = (n - C - k)^{-1} \sum_c \sum_i e_{Gic}^0 e_{Hic}^0, \tag{11a}$$

$$\tilde{\sigma}_{GG}^{11} = (n_G^+ - C - k)^{-1} \sum_c \sum_i (e_{Gic}^1)^2, \tag{11b}$$

$$\tilde{\sigma}_{GG}^{10} = (n_G^+ - C - k)^{-1} \sum_c \sum_i e_{Gic}^0 e_{Gic}^1, \tag{11c}$$

where n_G^+ is the sum of n_{cG}^+ over clusters and n is the total number of households. In (11b) and (11c) the summation is taken over all households that record unit values, while in (11a) it runs over all households. Note that eqs. (11b) and (11c) estimate only variances and covariances within goods, and that the covariances of the residuals *between* goods are assumed to be zero both within the unit-value equation and between the two equations.

In principle there is no difficulty in estimating the full matrices of inter-good covariances, and the appropriate formulae are the obvious extensions to (11b) and (11c). The formulae given below also apply to the more general case. However, in many applications there are relatively few households recording market purchases, and the estimation of off-diagonal elements for σ_{GH}^{10} and σ_{GH}^{11} could only be based on households that record purchases for *both* of the goods. Even then, each element of the estimated matrix would be formed from a different number of residuals, a fact that would have to be allowed for in calculating standard errors. The important covariances here are those between the budget shares and unit values for the same good. Expenditures and quantities are inevitably measured with error, so that when unit values are calculated by dividing one by the other, there will be generally be a correlation between the residuals in the budget share and unit value equations.

All of the first-stage estimators will be consistent as the sample size tends to infinity, even if the number of clusters increases at the same rate as the total sample size, as it would in the practically relevant case where cluster size is held constant. Of course the consistency comes at a price, that the parameters (including the variances) are the same for all clusters, so that the within-cluster information can be pooled over a large number of clusters.

The second stage of estimation begins by using the first-stage estimates to calculate the parts of mean cluster shares and unit values that are not accounted for by the first-stage variables. Define $\tilde{y}_{G.c}^0$ and $\tilde{y}_{G.c}^1$ by

$$\tilde{y}_{G.c}^0 = w_{G.c} - \tilde{\beta}_G^0 \ln x_{.c} - \tilde{\gamma}_G^0 \cdot z_{.c} = w_{G.c} - x_{.c} \cdot \tilde{\pi}_G^0, \tag{12a}$$

$$\tilde{y}_{G.c}^1 = \ln v_{G.c} - \tilde{\beta}_G^1 \ln x_{.c} - \tilde{\gamma}_G^1 \cdot z_{.c} = \ln v_{G.c} - x_{.c} \cdot \tilde{\pi}_G^1, \tag{12b}$$

where, in a slight abuse of notation, $x_{.c}$ is the vector of explanatory variables at the first stage and π_G^0 and π_G^1 are the parameters for the two equations. Define the matrix Q as the variance–covariance matrix across clusters of the theoretical magnitudes $y_{G.c}^0$, defined as above but using the true parameters β^0 and γ^0. S is the corresponding matrix for $y_{G.c}^1$ and R the covariance matrix, i.e.,

$$q_{GH} = \operatorname{cov}\left(y_{G.c}^0, y_{H.c}^0\right), \quad s_{GH} = \operatorname{cov}\left(y_{G.c}^1, y_{H.c}^1\right),$$
$$r_{GH} = \operatorname{cov}\left(y_{G.c}^0, y_{H.c}^1\right). \tag{13}$$

It is also convenient to have a matrix notation for the matrices of residual variances and covariances. Denote the population counterparts corresponding to (11a), (11b), and (11c) by Σ, Ω, and Γ, respectively. As defined above and as implemented in the computer code, the last two of these are diagonal matrices, but nothing in the theory prevents a more general interpretation.

From the population version of (12) and taking probability limits over all clusters,

$$S = \Psi M \Psi' + \Omega N_+^{-1}, \tag{14}$$

$$R = \Psi M \Theta' + \Gamma N^{-1}, \tag{15}$$

where M is the variance–covariance matrix of the unobservable price vector, $N_+^{-1} = \operatorname{plim} C^{-1}\sum_c D(n_c^+)^{-1}$, with $D(n_c^+)$ a diagonal matrix formed from the elements of n_{cG}^+, and N^{-1} is the corresponding matrix for the n_c's. Equating sample moments to their population counterparts, calculate the matrix $\tilde{B}$ according to

$$\tilde{B} = \left(\tilde{S} - \tilde{\Omega} T_+^{-1}\right)^{-1}\left(\tilde{R} - \tilde{\Gamma} T_A^{-1}\right), \tag{16}$$

where a tilde (~) denotes an estimate and the diagonal matrices T_A and T_+ are the sample counterparts of N and N_+ and are given by

$$T_A^{-1} = C^{-1}\sum_c \left\{D(n_c)\right\}^{-1}, \quad T_+^{-1} = C^{-1}\sum_c \left\{D(n_c^+)\right\}^{-1}, \tag{17}$$

and C is the total number of clusters in the sample. As the sample size goes to infinity with cluster sizes remaining fixed, $\tilde{B}$ will tend to its population counterpart, i.e.,

$$\operatorname{plim} \tilde{B} = B = (\Psi')^{-1}\Theta'. \tag{18}$$

It is not required that the cluster size becomes large; by pooling across clusters at the first stage, the first-stage parameters are consistent as the number of clusters increases. Similarly, the estimation of price effects rests entirely on the between-cluster variation, and the estimate of B will tend to its true value as the number of clusters grows large.

Estimates of B do not allow direct recovery of Ψ and Θ. However, eq. (10) together with (18) allows Θ to be calculated from

$$\Theta = B'\{I - D(\xi)B' + D(\xi)D(w)\}^{-1}. \tag{19}$$

The matrix of price elasticities E, from (6), is $\{D(w)\}^{-1}\Theta - \Psi$, so that, substituting,

$$E = \left\{D(w)^{-1}B' - I\right\}\{I - D(\xi)B' + D(\xi)D(w)\}^{-1}. \tag{20}$$

Estimates of Θ and E are calculated from (19) and (20) by replacing theoretical magnitudes with estimates from the first and second stages and by using the sample mean budget shares for the w vector.

The appendix derives variances and covariances for the parameters and for the elasticities, as well as test statistics for Slutsky symmetry. The remainder of this section reports a very limited Monte Carlo experiment with a 'stripped-down' version of the model. The model I have investigated is one in which there are neither cross-price nor quality effects; the focus is rather on the effects of the measurement error, particularly in the unit-value equation. The true model is formed from the following versions of (1) and (2):

$$w_{ic} = \alpha^0 + \beta^0 \ln x_{ic} + \theta \ln p_c + f_c + u_{ic}^0, \tag{21}$$

$$\ln v_{ic} = \ln p_c + u_{ic}^1. \tag{22}$$

There is only one good, and the absence of quality effects means that the unit value is simply a noisy measure of price. The parameter α^0 is set to zero, β^0 to 0.02, and θ, which is the parameter on which I shall focus, to 0.046; this value, together with the other assumptions to be made, generates a price elasticity of -0.67 at the mean of the budget share. Total expenditure x and the unobservable p are each generated by independent drawings from lognormal distributions, $\ln x$ with mean 4.6 and standard deviation 0.5, and $\ln p$ with mean zero and standard deviation 0.1. These values were drawn afresh at each experiment rather than held fixed in repeated samples; this appears to be the appropriate way to model repeated sample surveys from the same underlying population. I made no attempt to induce any cluster structure into the $\ln x$'s; all observations are independent drawings.

The fixed effects f_c are generated as $0.01(\ln x_c - 4.6) + 0.0159\varepsilon$, where 4.6 is the mean of $\ln x$, $\ln x_c$ is the cluster mean of $\ln x_{ic}$, and ε is drawn from N(0, 1). This procedure is chosen so as to generate a correlation of 0.3 between the fixed effects and $\ln x$. Finally, the two error terms u^0 and u^1 are independently drawn from normal distributions, each with mean zero and standard deviations of 0.0005 and 0.1, respectively. The independence of u^0 and u^1 corresponds to the case where errors of measurement in expenditures and prices are independent, so that there will be a *negative* correlation between measured quantity and measured unit value. While there is a good deal of arbitrariness in the choice of these parameters, they generate data that bear at least superficial resemblance to the results reported below.

The experiments were carried out as follows. Given a cluster size n_c, assumed to be the same for all clusters, and a number of clusters C, 500 'sample surveys' were generated according to the rules above and the estimate of θ calculated using eq. (16). No use was made of the absence of quality effects in the price equation, so that the unit values were regressed on $\ln x$ at the first stage as described in the text above. Based on the previous literature, two additional estimators were calculated. The first, referred to as the 'between-cluster' estimator, has a first stage that is identical to that for the 'correct' estimator, but at the second stage, the corrections for the estimated measurement errors are omitted. Referring to eq. (16), the estimate would be $\tilde{S}^{-1}\tilde{R}$. The basic idea is to ignore the within-cluster information, relying on the averaging within clusters to reduce the measurement error in prices; see Strauss (1982) for a similar argument in the context of farm-household behavior in Sierra Leone. The second, 'logarithmic', estimator follows the standard procedure of double-logarithmic regression; see for example Timmer and Alderman (1979) and Timmer (1981). In the experiments, $\ln q$ was calculated as $\ln w + \ln x - \ln p$, and then the individual household data were used to regress $\ln q$ on a constant, $\ln x$, and $\ln v$. To compare the estimated elasticity with the parameter θ, it was multiplied by the sample mean budget share and added to 1. A simple errors-in-variables analysis shows that, provided the true elasticity is greater than -1 as it is here, this estimator will be biased downward by the spurious negative correlation between $\ln q$ and $\ln v$.

The results are shown in graphical form in figs. 1, 2, and 3. In fig. 1 there are 100 clusters, each of size 2. (In the empirical results below, the situation is much better than this, and there are typically more than 2,000 clusters with 4–6 households in each.) The estimated densities are shown for each of the three estimators. They are calculated from the underlying 500 estimates by using a kernel density estimator with a Gaussian kernel and a bandwidth of $1.06 \cdot m^{-0.2} \cdot \min(sd, iqr/1.34\}$, where sd and iqr are the standard deviation and inter-quartile range of the estimates and $m = 500$ is the number of replications; see Silverman (1986, pp. 45–47). Unlike the other two estimators,

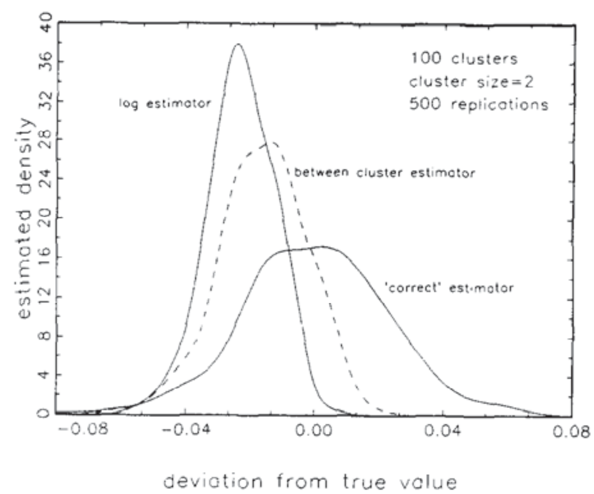

Fig. 1. Three estimators: smoothed densities.

the 'correct' estimator is correctly centered, but has noticeably greater spread. The between estimator is biased downward; averaging over clusters does not eliminate the measurement error in the unit values, so that the variance of 'prices' is overstated, thus producing an understatement when it is divided into the covariance of prices and budget shares. The expected downward bias in the logarithmic estimator is also apparent; in this case the elasticity would

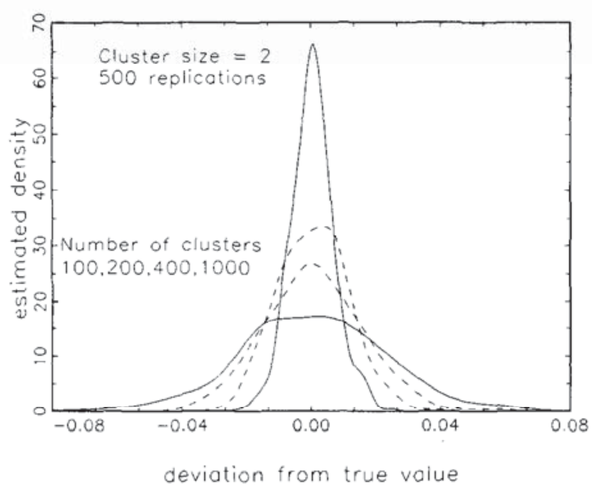

Fig. 2. Effects of numbers of clusters.

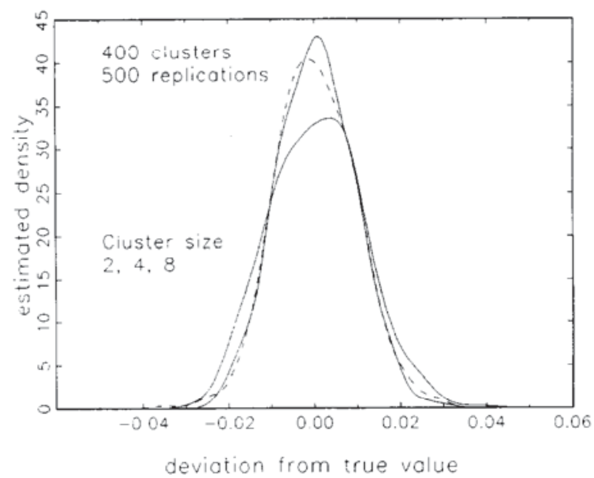

Fig. 3. Effects of cluster size.

be estimated as close to -1 instead of its true value (at the sample mean) of -0.67. The variance of the log estimator is the least because it uses the whole sample, not just the between-cluster data. That of the 'correct' estimator is relatively wide because of necessity to estimate the within-cluster error variances. The sampling variability of these variances causes a loss of precision but is required to correct the bias and inconsistency in the other two estimators. Not shown here is the bias of the total expenditure elasticity from the log estimator; by ignoring the fixed effects, the estimated expenditure elasticity is contaminated by the correlation between $\ln x$ and f, not to mention the measurement error in $\ln v$.

Figs. 2 and 3 show the effects of increasing, first, the number of clusters, and second, the number of households in each cluster. Fig. 2 shows the expected narrowing of the distribution as the number of clusters increases from 100 to 1000 with the cluster size at its worst possible value of 2. As theory would suggest, the standard deviations of the four densities are proportional to $C^{-0.5}$, and even with these relatively unfavorable sample sizes, the asymptotic standard errors derived in the appendix are an excellent approximation to the empirical standard deviations. For the four cases shown, with cluster size 2 and with 100, 200, 400, and 1000 clusters, the actual standard deviations with theoretical asymptotic standard errors in parentheses are, respectively, 0.0210 (0.0220), 0.0149, (0.0151), 0.0105 (0.0107), 0.0067 (0.0067). Fig. 3 shows that the larger cluster sizes are better than smaller ones, but the effect is not very marked. Again, this is what is predicted by the theory. Even with infinite numbers of households in each cluster, the estimates would not converge;

information about prices comes only from the cross-section of clusters. Again the asymptotic standard errors are a good guide to the actual results; for cluster sizes 4 and 8, the actual standard deviations (theoretical standard errors) are 0.0092 (0.0093) and 0.0086 (0.0088). Unless there is something special about these simulations (and they are the only ones that I have tried), the theory in the appendix would seem to provide useful formulae for standard errors, especially given that, in practice, sample sizes are likely to be much larger than those in the simulation.

3. Results from rural Java

The model described in section 2 was used to estimate a demand system of eleven foods using data from the 1981 SUSENAS household survey of Indonesia. Since spatial price variation is likely to be more marked in rural areas and since survey clusters are more widely spaced in the countryside, I report results only for rural households. In order to keep a sample that is relatively homogeneous, I further restrict attention to Java. Even so, the potential maximum sample is 14,487 households. Table 1 lists the eleven foods, together with their average budget shares, and the numbers of households providing information about each. Although households that record no expenditure are included in the analysis, whole clusters with zero expenditures are excluded because there is no way of estimating a price for them. There are 3,202 sample clusters in rural Java in all, and the third column of table 1 shows how many clusters record at least one household making a purchase for each of the goods. The first and second columns show how many households

Table 1

Commodities, sample sizes, and budget shares for Java, 1981.[a]

	Shares	Values	Clusters	Percent shares
Rice	12,914	9,245	2,804	24.53
Wheat	5,228	1,703	1,061	0.52
Maize	3,926	1,593	815	5.77
Cassava	6,441	2,539	1,343	1.39
Roots	5,716	2,021	1,185	0.60
Vegetables	14,419	14,115	3,181	5.57
Legumes	13,939	12,055	3,070	3.66
Fruit	10,114	4,652	2,124	1.88
Meat	4,928	1,526	1,001	2.07
Fresh fish	9,262	5,046	1,925	2.95
Dried fish	13,327	10,665	2,871	2.83

[a]First three columns are numbers of households in clusters with some households purchasing the good, numbers of households recording purchases, and numbers of such clusters. Final column shows budget shares for each good averaged over households in the first column.

Table 2

First-stage estimates: Quantity and quality effects.[a]

	β^0	$t(\beta^0)$	β^1	$t(\beta^1)$	ε	$t(\varepsilon)$
Rice	−0.1179	−56.7	0.0290	9.0	0.490	57.6
Wheat	0.0035	6.6	0.0993	1.1	1.567	23.4
Maize	−0.0526	−19.1	−0.0003	−0.0	0.088	3.3
Cassava	−0.0118	−14.9	0.0168	0.7	0.139	3.5
Roots	−0.0008	−1.7	0.1653	2.8	0.709	13.9
Vegetables	−0.0206	−33.7	−0.0402	−1.8	0.670	25.2
Legumes	−0.0039	−5.9	0.0422	4.6	0.850	42.3
Fruit	0.0086	12.4	0.0725	2.7	1.385	40.0
Meat	0.0287	18.2	0.0885	1.9	2.296	43.4
Fresh fish	0.0090	8.8	0.2232	8.5	1.082	35.0
Dried fish	−0.0105	−17.6	0.0639	5.7	0.566	25.5

[a]β^0 is the coefficient of the budget share and β^1 the coefficient of the unit value on the logarithm of household expenditure per capita. ε is the calculated total expenditure elasticity of quantity, β^1 is the total expenditure elasticity for quality.

are in these clusters and how many make purchases. Overall, there are just under five households per cluster. Even selecting out clusters where no household makes a purchase, the number of purchasers per cluster varies from 4.4 for vegetables to around 1.5 for meat. Expenditure on rice accounts for nearly a quarter of the budget, while the other foods account for another quarter between them.

Table 2 presents some of the results from the first-stage estimation. At this stage, cluster means are removed from all variables, and shares and logarithms of unit values regressed on the logarithm of household per capita expenditure, the logarithm of total household size, a set of demographic variables (the numbers of household members in each of thirteen age and sex categories as a ratio of household size), and nine educational dummies. The first four columns show, together with their t-values, the coefficients β^0 and β^1, which estimate the effects of the logarithm of total expenditure on the shares and the unit values. The last two columns show the total expenditure elasticities of quantity calculated according to eq. (4).

The quality elasticities β^1 relate to unit values, defined as expenditure per kilo, where the latter is calculated by adding weights across all goods in the group. Such a procedure makes sense for most of the goods, but is less than satisfactory for such categories as fresh fish, vegetables, or fruit, where I am (literally for fruit) adding apples and oranges. A more sophisticated calorie-based treatment would be an obvious next step. However, note that even if kilos of 'fruit' are of little interest on their own account, total weight is likely to be well-correlated with more satisfactory indicators of volume if the composition within the category is more or less constant across the sample. In any case, and with the exception of roots and fresh fish, the estimates in table 2 show very little response of unit value to total expenditure. For two goods, maize and vegetables, there are insignificant negative effects, while for the others, the elasticities are less than 10%. Even when the quality elasticity is relatively large, as in fresh fish, the effect is hardly dramatic. An estimated expenditure elasticity of fish *expenditure* of 1.31 is decomposed into an estimated quantity elasticity of fish of 1.08, with the difference of 0.22 representing the upgrading of quality at higher incomes. The quantity elasticities are ordered much as would be expected, with cassava and rice near the bottom, and meat, fresh fish, and wheat near the top.

Table 3 reports the estimates of variances and covariances from the first-stage estimates. The first column gives the diagonal of the matrix $\tilde{\Sigma}$, the residual variances from the share equation; the square roots of these numbers can be compared with the average budget shares from table 1 to give some idea of fit. The second column is the diagonal of Ω, while the third column, which is the most interesting, presents the covariances between the residuals in the two equations. The estimates of σ^{10} are important because they are informative about the importance of measurement error, and because their magnitude affects the size of the corrections that are made to what otherwise would be an OLS regression of cluster average shares on cluster average prices; see eq. (16) above. A natural starting point for discussion is the supposition that the errors of measurement in reported expenditures are orthogonal to the errors of measurement in reported quantities. If so, the error in the unit value will have a positive covariance with the error in the share, and that is what occurs in table 3 in all cases except for vegetables. The last may reflect a different reporting bias, or just the general inappropriateness of adding together the weights of different vegetables. The scale of σ^{10} will be determined by the

Table 3

First-stage variances, covariances, and cluster sizes.[a]

	σ^{00}	σ^{11}	σ^{10}	t_A	t_+
Rice	0.004915	0.00671	0.00072	3.85	2.31
Wheat	0.000138	0.50110	0.00122	4.28	1.28
Maize	0.002678	0.05991	0.00172	4.11	1.40
Cassava	0.000349	0.06175	0.00066	4.09	1.41
Roots	0.000100	0.30061	0.00023	4.14	1.33
Vegetables	0.000467	0.58771	−0.00180	3.74	3.59
Legumes	0.000530	0.08454	0.00041	3.76	2.99
Fruit	0.000450	0.21981	0.00119	4.08	1.59
Meat	0.001369	0.15164	0.00153	4.20	1.24
Fresh fish	0.000865	0.23510	0.00212	4.08	1.76
Dried fish	0.000413	0.10134	0.00082	3.92	2.68

[a] σ^{00}, σ^{11}, and σ^{10} are the residual variances of the share equation, the unit-value equation, and the covariance between them, respectively. t_A and t_+ are the elements of the matrices T_A and T_+ and are appropriately defined 'averages' of numbers of households per cluster, in total, and reporting purchases of the good.

Table 4

Cross-cluster variances and covariances and corrections.[a]

	cov(w, ln p)	var(ln p)	Ratio (1)	Ratio (2)	e_1	e_2
Rice	0.00203	0.0158	0.1284	0.1429	−0.48	−0.42
Wheat	0.00113	0.9545	0.0012	0.0015	−0.77	−0.71
Maize	0.00223	0.1499	0.0149	0.0169	−0.74	−0.71
Cassava	0.00117	0.1663	0.0071	0.0082	−0.49	−0.41
Roots	−0.00038	0.6486	−0.0006	−0.0010	−1.10	−1.17
Vegetables	−0.00184	0.5862	−0.0031	−0.0032	−1.06	−1.06
Legumes	−0.00032	0.1203	−0.0027	−0.0047	−1.07	−1.13
Fruit	0.00033	0.2424	0.0014	0.0004	−0.93	−0.98
Meat	−0.00010	0.2281	−0.0004	−0.0044	−1.02	−1.21
Fresh fish	0.00165	0.3311	0.0050	0.0057	−0.83	−0.81
Dried fish	0.00345	0.1404	0.0246	0.0316	−0.13	0.12

[a]The covariances and variances are evaluated across cluster means. Ratio (1) is the ratio of the covariance to the variance. Ratio (2) is the ratio of the covariance less σ^{10}/t_A to the variance less σ^{11}/t_+. e_1 and e_2 are the own-price elasticities calculated ignoring all cross-price and quality effects using ratio (1) and ratio (2), respectively.

budget share of the good, so that, once again, it is useful to deflate the estimates by the average budget share. Particularly noticeable is the very small figure for the most important commodity, rice, where σ^{10}/w is 0.0029, suggesting that measurement errors on expenditures have a standard error of only about 5%. For the other commodities, similar computations give much larger figures.

From an econometric point of view, the importance of allowing for measurement error is determined by the size of the corrections to S and R made by subtracting ΩT_A^{-1} and ΓT_+^{-1}. Table 4 is one way of assessing the extent to which measurement error affects the estimates. The first and second columns show the cross-cluster variance of mean unit value and its covariance with the budget share, where both quantities have been purged of the effects of the first-stage variables. The ratio of covariance to variance, ratio (1), is the OLS regression coefficient in the regression of mean budget share on mean unit value. If there were no quality effects, no cross-price effects, and no measurement error, this ratio, divided by the budget share, would be one plus the own price elasticity, e_1 in the table. As has already been seen, the quality effects are small in any case and, as will be seen, allowing for cross-price effects has only a minor effect on the own-price elasticities, so the simple estimates here are more useful than might appear. If the calculation is complicated only to the extent of allowing for measurement error, still ignoring cross-price and quality effects, the covariance is corrected by subtracting σ^{10}/t from table 3, the variance by subtracting σ^{11}/t_+, and the ratio and the elasticity recalculated. Hence ratio (2) and e_2 in the table.

Given the general uncertainty about price elasticities, the last two columns are remarkably close. Only for dried fish, and to a lesser extent for meat, does the correction make a real difference. For some commodities, for example rice, the estimated size of the measurement error is small relative to the substantive variances and covariances so that ratio (1) and ratio (2) are similar. This finding replicates earlier work on the demand for rice in Indonesia by Case (1988). For other goods, allowing for measurement error makes a large proportional difference to the ratio of covariance to variance, but since the ratio itself is small, the estimated elasticity remains close to minus one. In either case, similar results would have been obtained by simply ignoring the measurement error, provided, of course, that the regression is one of average cluster demand on average cluster price. However, it is important to note that this result provides no support for a regression, at the household level, of log quantity on log price. Such a regression is restricted to households that make positive purchases and the results on the restricted sample are likely to be (and in this case are) quite different. Even simple regressions of budget shares on $\ln x$ have quite different coefficients depending on whether or not zeros are excluded, particularly for those commodities where there is a large number of households not purchasing. Further, by working at the household level rather than with cluster means, there is no averaging to reduce the effects of measurement error relative to the effects of genuine price variation.

The matrix of own- and cross-price elasticities is presented in tables 5a and 5b. All of the estimated own-price elasticities are negative, as they should be. Note however that the 'default' own-price elasticity is not zero but minus one, a value that is attained when B and β^1 are zero; see eq. (20). The own-price elasticities are also close to the preliminary estimates in table 4; allowing for cross-price effects is not very important for measuring own-price elasticities, at least in these data. For four commodities (vegetables, legumes, fruit, and meat) the price elasticity is not significantly different from this default value – a phenomenon that may reflect problems with defining quantities. For all goods except meat and for all goods with price elasticities significantly different from minus one, the price elasticities are estimated to lie between minus one and zero. There are no goods with very large estimated own-price elasticities, and there is some tendency for the goods that have the lowest total expenditure elasticities (dried fish, cassava, rice) to have absolutely low price elasticities – something that might be expected for goods that are are genuinely 'necessary'.

The important rice price elasticity is estimated to be -0.42, which can be compared with the figures of -0.55 or -0.62 (depending on detailed assumptions) estimated from a subsample of 5218 households in Java from the same survey by van de Walle (1988). Van de Walle used a log-log specification, and the difference in the estimates is in the direction that the theory would predict. However, the divergence is not very large; the vast majority of households

Table 5a

Own- and cross-price elasticities for Java, 1981.[a]

	1 Rice	2 Wheat	3 Maize	4 Cassava	5 Roots	6 Vegetables
1 Rice	−0.424	−0.005	−0.032	0.088	0.025	−0.057
	5.1	0.3	0.9	3.1	1.8	3.2
2 Wheat	−0.461	−0.692	0.011	−0.000	−0.188	0.284
	1.6	13.8	0.1	0.0	3.8	4.6
3 Maize	1.245	0.090	−0.822	−0.256	−0.100	0.602
	4.9	2.1	7.5	3.0	2.4	9.5
4 Cassava	0.151	0.078	0.186	−0.325	−0.003	0.121
	0.5	1.4	1.4	2.8	0.1	1.7
5 Roots	0.795	0.125	0.150	0.084	−0.953	0.080
	3.9	3.6	1.9	1.2	22.0	1.8
6 Vegetables	−0.047	−0.057	−0.100	0.101	0.030	−1.113
	0.4	3.1	2.3	2.7	1.6	28.5
7 Legumes	0.108	−0.038	−0.200	−0.050	0.034	−0.118
	0.8	1.6	3.5	1.0	1.4	3.8
8 Fruit	0.354	−0.013	−0.003	0.171	−0.068	0.056
	2.1	0.4	0.0	2.9	2.3	1.5
9 Meat	−0.190	−0.075	−0.060	−0.173	−0.158	0.176
	0.8	1.7	0.6	2.0	3.6	3.2
10 Fresh fish	0.399	0.028	0.244	−0.017	−0.099	0.316
	2.0	0.8	3.0	0.2	2.8	7.0
11 Dried fish	0.391	0.029	0.151	0.057	−0.057	0.379
	2.0	0.8	1.9	0.8	1.7	8.4

[a] The column is the good whose price is changing and the row is the good affected. Hence, a increase in the price of vegetables is estimated to increase the consumption of wheat by 0.284. The figures below the elasticities are (absolute) asymptotic t-values.

consume rice, and van de Walle argues that the rice price may be relatively well measured. Timmer and Alderman (1979) and Timmer (1981) also report price elasticities for rice and cassava from the 1976 survey. Again they use a double logarithmic formulation, but apply the model not to the micro data but to cell-means of income classes by province, sector, and time period. They estimate different elasticities for different income groups and find figures that are numerically very much larger than those reported here. Most income groups have rice price elasticities below -1, and the average is -1.1, while for fresh cassava, the average price elasticity is -0.8, as opposed to -0.33 in table 5a. Timmer and Alderman also report a cross-price elasticity of cassava with respect to the rice price of 0.77 (0.15 here), but fail to find a significant cross-price effect from cassava to rice (0.09 here). There are many possible reasons for these discrepancies, but again the double logarithmic quantity on unit value specification must be a prime suspect. I also find it implausible that such basic staples should display such high price elasticities, if only because there are few obvious substitutes. However, the opposite position has also

Table 5b
Own- and cross-price elasticities for Java, 1981.[a]

	7 Legumes	8 Fruit	9 Meat	10 Fresh fish	11 Dried fish
1 Rice	0.065	0.130	−0.064	0.053	0.213
	2.0	2.9	1.6	2.1	6.3
2 Wheat	−0.127	−0.211	−0.226	0.089	−0.128
	1.1	1.3	1.6	1.0	1.1
3 Maize	0.006	−0.247	0.167	−0.444	−0.111
	0.1	1.8	1.4	5.6	1.1
4 Cassava	0.013	−0.461	0.184	0.054	−0.671
	0.1	2.5	1.1	0.5	5.0
5 Roots	−0.122	−0.072	0.080	0.181	−0.253
	1.5	0.6	0.8	2.9	3.1
6 Vegetables	−0.081	−0.064	−0.064	0.050	−0.173
	1.9	1.1	1.2	1.5	4.0
7 Legumes	−0.954	−0.097	0.202	−0.084	−0.134
	0.8	1.3	3.0	1.9	2.4
8 Fruit	−0.088	−0.953	−0.092	0.071	−0.070
	1.3	0.5	1.1	1.4	1.0
9 Meat	0.030	−0.376	−1.091	−0.036	0.196
	0.3	2.7	0.7	0.5	1.9
10 Fresh fish	0.026	−0.009	0.047	−0.762	0.259
	0.3	0.1	0.5	3.8	3.2
11 Dried fish	−0.023	−0.105	−0.057	0.667	−0.239
	0.3	1.0	0.6	10.4	9.4

[a]See footnote to table 5a.

been argued: that poor, near-subsistence consumers will substitute between minimum-cost, calorie-based diets as relative prices change.

There are a considerable number of estimates of cross-price elasticities that are significantly different from zero. It is not difficult to invent 'explanations' to account for almost any observed pattern of responses, but undoubtedly some of the figures make a good deal more sense than others. For example, the elasticities are gross elasticities, with income effects included, so that it would be reasonable to expect many of the numbers in the first column to reflect the large negative income effects generated by increases in the price of rice. Yet most of these estimates are *positive*, and there are substantial proportional increases in the demand for maize, roots, fruit, and fish when the price of rice rises. For this to make sense, the other omitted categories must show an overall complementarity with rice. Some other cross-price effects, for example the strong substitutability between fresh and dried fish, look a good deal more satisfactory. Table 6 shows the deviations from symmetry evaluated according to eq. (39), together with their absolute t-values. Sixteen out of the fifty-five cross-terms are significantly different from zero, and the overall Wald statistic

Table 6
Deviations from symmetry.[a]

	2	3	4	5	6	7	8	9	10	11
1	0.1	**7.4**	**−2.1**	−0.1	1.4	−0.9	−2.1	2.1	0.3	**−4.1**
	0.0	4.6	2.5	0.3	1.9	0.9	1.8	1.9	0.4	4.0
2		0.5	0.1	**0.2**	**−0.5**	−0.1	0.1	−0.0	0.3	0.1
		1.9	1.0	5.1	4.6	0.8	0.8	0.3	0.2	1.2
3			**1.7**	**0.7**	**−3.8**	−0.6	1.6	−0.8	**3.4**	1.1
			3.3	2.8	8.6	1.0	2.0	1.1	6.6	1.8
4				0.1	0.4	−0.2	**1.0**	−0.6	−0.1	**1.1**
				0.7	1.9	0.6	3.6	1.9	0.3	4.1
5					0.1	0.2	−0.1	**−0.3**	**−0.4**	−0.0
					1.1	2.0	0.9	3.3	3.6	0.1
6						0.1	0.5	**0.9**	**0.7**	**2.0**
						0.2	1.6	3.0	3.2	7.4
7							0.2	−0.6	0.4	0.4
							0.7	1.7	1.5	1.3
8								−0.6	−0.2	−0.2
								1.7	0.5	0.6
9									0.1	−0.7
									0.4	1.9
10										**1.1**
										3.6

[a]The upper entry in row i, column j is (100 times) the quantity $w_j e_{ji} - w_i e_{ij} + w_j w_i (e_j - e_i)$, which should be zero under Slutsky symmetry. The lower figures are absolute asymptotic t-values. Deviations with t-values greater than 2 are shown in bold.

for the null of symmetry is 415.8, a number that is well in excess of conventional critical values of a χ^2, even with 55 degrees of freedom. In Deaton (1987), I reported a favorable symmetry test for the Cote d'Ivoire; however, in that case there were only 5 goods and less than 200 clusters, as opposed to 11 goods and 2000 clusters in the current example. With such large samples, standard hypothesis tests will always tend to reject if no attempt is made to trade off Type I and Type II errors, and a case can be made for using a Bayesian procedure, such as that advocated by Schwartz (1978). This would reject the null only for test statistics greater than 55 times the log of the sample size, or 418.0. While such a procedure is not universally accepted, it is perhaps reasonable to conclude that the evidence is not overwhelmingly against the symmetry hypothesis.

4. Conclusions

This paper has proposed a method for using large-scale household survey data for the estimation of a system of demand equations, making use of spatial variation in price to identify and estimate a matrix of own- and cross-price elasticities. Compared with earlier formulations [Deaton (1987, 1988)], the model presented here contains a more satisfactory specification of the demand functions, as well as a derivation of the rather complex formulae for variances and covariances of the estimates. The admittedly limited Monte Carlo evidence presented in section 2 does not suggest any practical problems in using the procedure or the formulae for the standard errors. Large-scale sample surveys provide the sort of data where asymptotic theory is likely to provide a very good approximation.

Estimates are presented for an eleven-commodity system for Indonesia. To the extent that it is possible to judge, the parameter estimates are plausible. For rice and cassava, the own-price elasticities are a good deal smaller (i.e., closer to zero) than some previous estimates, but the direction of the discrepancy is as predicted by the theory developed here. If judged by a large-sample posterior odds test, the estimates are consistent with Slutsky symmetry, but the hypothesis would be rejected by a strict classical statistician (as would most sharp hypotheses on a sample of more than 14,000 observations).

While the model appears to work well in the application, there are a number of unresolved issues that ought to be noted. The model is very close to being exactly identified, and so it is difficult to construct the sort of cross-checks that would lend it greater conviction. Plausibility of demand elasticities is not in itself a very powerful test. It would be extremely desirable to have data with direct measures of market prices against which this method could be compared. The model would also be improved by allowing a more general functional form for the Engel curve in eq. (1). Given that zero observations are included, there are even less grounds than usual for assuming linearity. An ideal, but technically difficult solution would be a semiparametric form in which the Engel effects are dealt with nonparametrically. Although much remains to be done, the results reported here would nevertheless support the cautious use of the procedure in practical applications.

Appendix

This appendix describes the procedures for deriving standard errors for the parameters B, as well as the derived quantities Θ and E. As pointed out by a referee, the estimator in section 2 works by equating sample with theoretical moments, and so belongs to the generalized methods of moments class of estimators. The general theory of this class has been worked out by Hansen (1982), so that one approach to deriving standard errors would be to apply Hansen's formulae. The sequential nature of the procedure can be handled by the results given by Newey (1984). The GMM method can be used to explore efficiency considerations, as well as to check the extent to which it is possible to relax the assumptions of homoskedasticity and normality which are made below. However, the algebra involved is still extraordinarily heavy, and following this route all the way would be a research project in its own right. Given computer software that would handle symbolic matrix differentiation, it would be a useful exercise to redo the calculations below using the GMM method. Here, I proceed directly under the assumption that the error terms in (1) and (2) are i.i.d. normal.

I start with the variance–covariance matrix of the estimate of the matrix B; this is the most straightforward to deal with, and the variances of the elements of E are derived from it. The basic procedure is the 'delta method' [see for example Fuller (1987, p. 108)]. Expand $\tilde{B}$ from eq. (16) around the true value B, so that

$$(\tilde{B} - B) = (S - \Omega T_+^{-1})^{-1} \times \{(\tilde{R} - R) - (\tilde{\Gamma} - \Gamma)T_A^{-1} - (\tilde{S} - S)B + (\tilde{\Omega} - \Omega)T_+^{-1}B\}. \tag{A.1}$$

The derivations are simplified by adopting the following notation from Deaton (1987). Define the following matrices:

$$H = \begin{pmatrix} Q & R' \\ R & S \end{pmatrix}, \qquad \Lambda = \begin{pmatrix} \Sigma & \Gamma' \\ \Gamma & \Omega \end{pmatrix}, \tag{A.2}$$

together with $J = (O_N | I_N)$, an $N \times 2N$ matrix of ones and zeros, and $P' = (I_N | -B')$. Then (A.1) becomes

$$(\tilde{B} - B) = A^{-1}J\{(\tilde{H} - H) - (\tilde{\Lambda} - \Lambda)T^{-1}\}P, \tag{A.3}$$

where $A = (S - \Omega T_+^{-1})$ and T is a $2N \times 2N$ diagonal matrix with T_A on the first N elements of the diagonal and T_+ on the second N elements. Eq. (A.3) traces the sources of variance in $\tilde{B}$ back to the estimation uncertainties in $\tilde{H}$ and in $\tilde{\Lambda}$. The latter comes entirely from the first-stage estimation and can be dealt with straightforwardly using the fact that $\tilde{\Lambda}$ is a Wishart matrix. It is the estimation error in $\tilde{H}$ that is complicated because there are two independent sources: (a) the estimation error that comes from the estimation of parameters at the first stage and (b) the inherent uncertainty that comes from estimating population from sample covariance matrices, an uncertainty that would be present even if the first-stage parameters were known. In Deaton (1987), albeit using a different functional form, I incorrectly assumed that source (a) could be asymptotically ignored. The numerical results in that paper are barely affected by the error, but in general both sources of variance need to be taken into account, even in large samples.

To decompose further the estimation error $(\tilde{H}-H)$, rewrite (12) in the form

$$\tilde{Y}^0 = Y^0 - X(\tilde{\Pi}^0 - \Pi^0), \tag{A.4a}$$

$$\tilde{Y}^1 = Y^1 - X(\tilde{\Pi}^1 - \Pi^1), \tag{A.4b}$$

where Y^0 and Y^1 are the $C \times N$ matrices of cluster means of budget shares and log unit values, each with the grand mean removed, X is the corresponding $C \times K$ matrix of cluster means of first-stage explanatory variables, again with the grand mean removed, and Π^0 and Π^1 are $K \times N$ matrices of first-stage parameters. By aligning the matrices to define $Y = (Y^0|Y^1)$, a $C \times 2N$ matrix, and similarly for $\Pi = (\Pi^0|\Pi^1)$, (A.4a) and (A.4b) can be combined to give

$$\tilde{Y} = Y - X(\tilde{\Pi} - \Pi). \tag{A.5}$$

The matrix $\tilde{H}$ is then simply $C^{-1}\tilde{Y}'\tilde{Y}$, and I denote the corresponding estimate with known first-stage parameters as $\hat{H}$, i.e., $\hat{H} = C^{-1}Y'Y$. Write M for the covariance matrix $C^{-1}X'Y$, so that, to the first order,

$$\tilde{H} = \hat{H} - M'(\tilde{\Pi} - \Pi) - (\tilde{\Pi} - \Pi)'M. \tag{A.6}$$

Substituting in (A.3), we have

$$(\tilde{B} - B) = A^{-1}J\{(\hat{H} - H) - M'(\tilde{\Pi} - \Pi) - (\tilde{\Pi} - \Pi)'M - (\tilde{\Lambda} - \Lambda)T^{-1}\}P, \tag{A.7}$$

which isolates the three sources of variance in the estimate of B. The first expression in the braces, $\hat{H} - H$, comes from the second stage between cluster residuals and has a standard Wishart form. The variability in the estimates of Π and Λ comes from the first-stage within-cluster residuals. Since the within estimates are orthogonal to the between estimates, there is no covariance between the first term and the last three. Further, since the Π parameters are obtained by linear regression at the first stage, their estimates are asymptotically independent of the estimates of the variances and covariances of the residuals, and thus of the estimates of Λ. Note finally that no allowance need be made for the variability of M that comes from the variability of Y, since the effects through (A.6) are of the second order. Given this, I have economized on notation by not replacing M in (A.7) by its population counterpart.

In 'vec' notation (A.7) takes the form

$$\begin{aligned}\operatorname{vec}(\tilde{B}-B) &= (P' \otimes A^{-1}J)\big[\operatorname{vec}(\tilde{H}-H) - \operatorname{vec}\{M'(\tilde{\Pi}-\Pi)\} \\ &\quad - \operatorname{vec}\{(\tilde{\Pi}-\Pi)'\tilde{M}\} - \operatorname{vec}\{(\tilde{\Lambda}-\Lambda)T^{-1}\}\big] \\ &= (P' \otimes A^{-1}J)\big[\operatorname{vec}(\tilde{H}-H) \\ &\quad - (I+K)(I \otimes M')\operatorname{vec}(\tilde{\Pi}-\Pi)\big] \\ &\quad - (P'T^{-1} \otimes A^{-1}J)\operatorname{vec}(\tilde{\Lambda}-\Lambda),\end{aligned} \tag{A.8}$$

where K is the $2N^2 \times 2N^2$ commutation matrix with the property that $K \operatorname{vec}(A) = K \operatorname{vec}(A')$ for arbitrary conformable matrix A; see Magnus and Neudecker (1986). Given that both $\tilde{H}$ and $\tilde{\Lambda}$ are Wishart matrices, we have

$$\mathrm{E}\big[\operatorname{vec}(\tilde{H}-H)\operatorname{vec}(\tilde{H}-H)'\big] = C^{-1}(H \otimes H)(I+K), \tag{A.9}$$

$$\mathrm{E}\big[\operatorname{vec}(\tilde{\Lambda}-\Lambda)\operatorname{vec}(\tilde{\Lambda}-\Lambda)'\big] = (n-C-k)^{-1}(\Lambda \otimes \Lambda)(I+K), \tag{A.10}$$

while, since Π is estimated at the first stage from the standard multivariate regression model,

$$\mathrm{E}\big[\operatorname{vec}(\tilde{\Pi}-\Pi)\operatorname{vec}(\tilde{\Pi}-\Pi)'\big] = \Lambda \otimes (W'W)^{-1}, \tag{A.11}$$

where W is the matrix of explanatory variables for the within regression at the first stage. Note that the matrices W and X relate to the same variables, but that X consists of deviations of cluster means from the grand mean, while W consists of deviations of individual observations from their cluster means.

Combining eqs. (A.8) through (A.11) yields

$$V\big[\operatorname{vec}(\tilde{B})\big] = V_1 + V_2 + V_3, \tag{A.12}$$

where

$$V_1 = C^{-1}(P'HP \otimes A^{-1}JHJ'A^{-1}) + C^{-1}(P'HJ'A^{-1} \otimes A^{-1}JHP)K,$$

$$\begin{aligned}V_2 &= (n-C-k)^{-1}(P'T^{-1}\Lambda T^{-1}P \otimes A^{-1}J\Lambda J'A^{-1}) \\ &\quad + (n-C-k)^{-1}(P'T^{-1}\Lambda J'A^{-1} \otimes A^{-1}J\Lambda T^{-1}P)K,\end{aligned}$$

$$V_3 = (P' \otimes A^{-1}J)(I+K)\big(\Lambda \otimes M'(W'W)^{-1}M\big)(I+K)(P \otimes J'A^{-1}).$$

Although (A.12) is the simplest way to write the variance–covariance matrix, the expression for V_3 involves a larger than necessary matrix, since the matrix in the center is $4N^2 \times 4N^2$, which can be large enough to cause storage difficulties. The following equivalent expression uses smaller matrices:

$$\begin{aligned} V_3 = {} & P'\Lambda P \otimes A^{-1}JM'(W'W)^{-1}MJ'A^{-1} \\ & + P'M'(W'W)^{-1}MP \otimes A^{-1}J\Lambda J'A^{-1} \\ & + \left(P'M'(W'W)^{-1}MJ'A^{-1} \otimes A^{-1}J\Lambda P\right)K \\ & + \left(P'\Lambda J'A^{-1} \otimes A^{-1}JM'(W'W)^{-1}MP\right)K. \end{aligned} \tag{A.13}$$

Note finally that there is some ambiguity in implementing V_3, because (A.11) does not reflect the fact that the unit-value equations are estimated on a smaller sample than are the share equations. There are several ways of fixing this that give the right answer in large samples. Here I take $W'W$ as the matrix from the share regressions and scale Ω, the bottom right-hand matrix of Λ, by n/n^+, where n^+ is the average over goods of the numbers of observations entering the unit-value equations. Since the first-stage parameters are based on a sample that is typically ten times larger than the sample at the second stage, both V_2 and V_3 usually make very small contribution to the total variance.

Although the above formulae give variances and covariances for all the parameters that are estimated, in most applications interest will focus not on the B parameters but on the estimates of elasticities that are derived from them. In principle, it is straightforward to use the delta method to estimate variances and covariances of functions of the elements of B. In practice, the algebra is extremely tedious. An outline is included here for completeness and to support the computer code. Start with the total expenditure elasticities, ε_G, which depend only on the first-stage parameters β_G^0 and β_G^1; see eq. (4). Write the total differential of (4) in the form $\mathrm{d}\varepsilon = \Phi_1\,\mathrm{d}\Pi' e = (e' \otimes \Phi_1)K \operatorname{vec} \mathrm{d}\Pi$, where e is a (basis) vector of zeros with a one in the position occupied by $\ln x$ in the matrix of first-stage variables, where Π, as before, is the $K \times 2N$ matrix of first-stage parameters and $\Phi_1 = [I| - D(w)^{-1}]$ is the $N \times 2N$ Jacobian matrix of partial derivatives. It then follows that

$$V(\tilde{\varepsilon}) = (e' \otimes \Phi_1)K\left(\Lambda \otimes (W'W)^{-1}\right)K(e \otimes \Phi_1') = \omega\Phi_1\Lambda\Phi_1', \tag{A.14}$$

where $\omega = e'(W'W)^{-1}e$ is the element in the diagonal of $(W'W)^{-1}$ corresponding to $\ln x$.

Turning next to the price elasticities E, transpose eq. (20) and take total differentials to give

$$\mathrm{d}E' = G\cdot \mathrm{d}B\left[D(w)^{-1} + D(\xi)E'\right] + G[B - D(w)]\,\mathrm{d}D(\xi)\cdot E', \tag{A.15}$$

where $G = \{I - B\xi + D(\xi)D(w)\}^{-1}$. Hence, in vec notation,

$$\begin{aligned}\operatorname{vec}(\mathrm{d}E') &= \left\{\left(D(w)^{-1} + ED(\xi)\right)\otimes G\right\}\operatorname{vec}\mathrm{d}B \\ &\quad + \{E\otimes G(B - D(w))\}\operatorname{vec}D(\mathrm{d}\xi).\end{aligned} \tag{A.16}$$

To evaluate the second term on the right-hand side, note that ξ_G given by (9), like ε_G above, is a function only of the first-stage parameters β_G^0 and β_G^1. Define the $N\times 2N$ matrix $\Phi_2 = [D(\rho^0)|D(\rho^1)\}$, where $\rho_G^0 = \partial\xi_G/\partial\beta_G^0$ and $\rho_G^1 = \partial\xi_G/\partial\beta_G^1$ evaluated from (9), so that in parallel with the discussion of ε above, we have $\mathrm{d}\xi = \Phi_2\,\mathrm{d}\Pi' e = (e'\otimes\Phi_2)K\operatorname{vec}\mathrm{d}\Pi$. The 'diagonalization' matrix L, an $N^2\times N$ matrix of ones and zeros, is defined by its property $L\xi = \operatorname{vec}\{D(\xi)\}$, for any N vector ξ. It can be used to express $\operatorname{vec}D(\mathrm{d}\xi)$ as $L\,\mathrm{d}\xi = L(e'\otimes\Phi_2)K\operatorname{vec}\mathrm{d}\Pi$, so that, combining (A.8) and (A.16),

$$\operatorname{vec}(\tilde{E}' - E') = F_1\operatorname{vec}(\tilde{H} - H) + F_2\operatorname{vec}(\tilde{\Lambda} - \Lambda) + F_3\operatorname{vec}(\tilde{\Pi} - \Pi), \tag{A.17}$$

where

$$F_1 = \left[D(w)^{-1} + ED(\xi)\right]P'\otimes GA^{-1}J,$$

$$F_2 = -\left[D(w)^{-1} + ED(\xi)\right]P'T^{-1}\otimes GA^{-1}J,$$

$$F_3 = [E\otimes G\{B - D(w)\}]L(e'\otimes\Phi_2)K - F_1(I+K)(I\otimes M'),$$

so that, given the asymptotic independence of the estimates of H, Λ, and Π, a formula for the variance–covariance matrix can be derived. Unfortunately, this straightforward procedure leads to large matrices, including, for example, the last commutation matrix above, which has $4N^2K^2$ elements. After some manipulation, it is possible to derive the following expressions which contain matrices no larger than $\max(4N^2, 2NK)$:

$$V(\operatorname{vec}\tilde{E}') = C^{-1}V_a + (n-k-C)^{-1}V_b + V_{11} + V_{12} + V_{12}' + V_{22}, \tag{A.18}$$

where

$$V_a(P,H) = \left[D(w)^{-1} + ED(\xi)\right] P'HP \left[D(w)^{-1} + D(\xi)E'\right]$$

$$\otimes GA^{-1}JHJ'A^{-1}G',$$

$$V_b = V_a(PT^{-1}, \Lambda),$$

$$V_{11} = \omega\left[E \otimes G\{B - D(w)\}\right] L\Phi_2\Lambda\Phi_2'L'\left[E' \otimes \{B' - D(w)\}G'\right],$$

$$V_{12} = -\left[E \otimes G\{B - D(w)\}\right]$$

$$\times L\left[e'(W'W)^{-1}MP\{D(w)^{-1} + D(\xi)E'\} \otimes \Phi_2\Lambda J'A^{-1}G'\right]$$

$$-\left[E \otimes G\{B - D(w)\}\right]$$

$$\times L\left[e'(W'W)^{-1}MJ'A^{-1}G' \otimes \Phi_2\Lambda P\{D(w)^{-1} + D(\xi)E'\}\right]K,$$

$$V_{22} = \left[\{D(w)^{-1} + ED(\xi)\} \otimes G\right] V_3\left[\{D(w)^{-1} + D(\xi)E'\} \otimes G'\right].$$

In the consumer-demand literature, there has frequently been an interest in examining the matrix of cross-price effects for evidence of Slutsky symmetry. In the notation used here, symmetry is satisfied at the budget shares w if the following condition holds:

$$\Delta = D(w)E + D(w)\varepsilon w' - E'D(w) - we'D(w) = 0. \tag{A.19}$$

In vec notation, this can be written

$$\begin{aligned}\delta &= \operatorname{vec}(\Delta)\\ &= (K - I)\{I \otimes D(w)\}\operatorname{vec} E' + (K - I)\{D(w) \otimes w\}\varepsilon \qquad \text{(A.20)}\\ &= 0.\end{aligned}$$

An estimate of Δ can be calculated from the estimates of E and ε, and a Wald test constructed given a variance–covariance matrix for Δ. Note that Δ is the difference between a matrix and its transpose, so it has a zero diagonal and an upper right triangle that is minus its lower right triangle. In consequence, only the elements of δ corresponding to the bottom left-hand triangle (below the diagonal) of Δ need be used for inference. Denoting these by δ^* and the corresponding variance–covariance matrix by $V(\delta^*)$, the Wald test is $\delta^{*\prime}\{V(\delta^*)\}^{-1}\delta^*$. The matrix $V(\delta^*)$ is selected from the elements of $V(\delta)$ obtained from using (A.20) above. Note that the variance–covariance matrices for vec $\tilde{E}'$ and for $\tilde{\varepsilon}$ have already been obtained, so that the only further requirement is for the covariance cov(vec $\tilde{E}'$, $\tilde{\varepsilon}'$}. Referring to (A.17) gives

$$\begin{aligned} C_v &= E\left[\operatorname{vec}(\tilde{E}' - E')(\tilde{\varepsilon} - \varepsilon)'\right] \\ &= F_3\left[\Lambda \otimes (W'W)^{-1}\right] K(e \otimes \Phi_1'), \end{aligned} \tag{A.21}$$

which, after substitution from (A.17) and rearrangement, gives

$$\begin{aligned} C_v = {} & \left[E \otimes G\{B - D(w)\}\right] L\Phi_2\Lambda\Phi_1'\omega \\ & - \left\{D(w)^{-1} + ED(\xi)\right\} P'M'(W'W)^{-1}e \otimes GA^{-1}J\Lambda\Phi_1' \\ & - \left\{D(w)^{-1} + ED(\xi)\right\} P'\Lambda\Phi_1' \otimes GA^{-1}JM'(W'W)^{-1}e. \end{aligned} \tag{A.22}$$

The variance–covariance matrix of $\tilde{\delta}$ is then constructed from the outer product of the estimate of (A.20), substituting from (A.14) and (A.18) for the two variances and from (A.22) and its transpose for the two covariances.

References

Case, Anne C., 1988, Analyzing spatial patterns in developing country data, Ph.D. thesis (Princeton University, Princeton, NJ).

Deaton, Angus S., 1986, Quality, quantity, and spatial variation of price (Princeton University, Princeton, NJ) processed.

Deaton, Angus S., 1987, Estimation of own- and cross-price elasticities from survey data, Journal of Econometrics 36, 7–30.

Deaton, Angus S., 1988, Quality, quantity, and spatial variation of price, American Economic Review 78, 418–430.

Deaton, Angus S. and John Muellbauer, 1980, An almost ideal demand system, American Economic Review 70, 312–326.

Fuller, Wayne A., 1987, Measurement error models (Wiley, New York, NY).

Hansen, Lars P., 1982, Large sample properties of generalized methods of moments estimators, Econometrica 50, 1029–1054.

Laraki, Karim, 1988, The nutritional, welfare, and budgetary effects of price reform in developing countries: Food subsidies in Morocco (Welfare and Human Resources Unit, The World Bank, Washington, DC) processed.

Magnus, Jan R. and H. Neudecker, 1986, Symmetry, 0–1 matrices and Jacobians: A review, Econometric Theory 2, 157–190.

Newey, Whitney K., 1984, A method of moments interpretation of sequential estimators, Economics Letters 14, 201–206.

Schwartz, Gideon, 1978, Estimating the dimension of a model, Annals of Statistics 6, 461–464.

Silverman, Bernard, 1986, Density estimation for statistics and data analysis (Methuen, New York, NY).

Strauss, John, 1982, Determinants of food consumption in rural Sierra Leone, Journal of Development Economics 11, 327–353.

Timmer, C. Peter, 1981, Is there 'curvature' in the Slutsky matrix?, Review of Economics and Statistics 63, 395–402.

Timmer, C. Peter and Harold Alderman, 1979, Estimating consumption parameters for food policy analysis, American Journal of Agricultural Economics 61, 982–987.

Van de Walle, Dominique, 1988, On the use of the Susenas for modelling consumer behavior, Bulletin of Indonesian Economic Studies 24, 107–122.