Department of Agricultural and Resource Economics
University of California, Davis

Quantity Versus Shares in Estimating Demand Systems

by

Quirino Paris and Francesco Caracciolo

Working Paper No. 12-003

2012

Giannini Foundation of Agricultural Economics

Quantity Versus Shares in Estimating Demand Systems

Quirino Paris*,† and Francesco Caracciolo§

† University of California, Davis, USA; § University of Naples, Italy

Abstract

This paper considers the estimation and testing of demand systems when the number of sample goods is smaller than the number of commodity choices available to consumers. In this case, the demand system is incomplete. The large majority of papers that appeared in the literature specifies and estimates a demand system in share format even when the system may be incomplete. The criterion for deciding whether a share format is admissible without loss of information is a test of the adding-up condition. This test, however, requires the estimation of a demand system in quantity format.

Keywords: demand systems, quantity format, share format, adding up

JEL classification: C01, C31, C33, D12

*Corresponding author: Department of Agricultural and Resource Economics, University of California, Davis, One Shields Ave, Davis CA, 95616, USA. Email: paris@primal.ucdavis.edu

1. Introduction

The objective of this paper is to discuss the advantages and limitations of estimating systems of demand functions that are specified in expenditure-share format and in original quantity format.

To summarize the discussion elaborated further on, the advantages of a share format may be listed as saving degrees of freedom and mitigating error heteroskedasticity. The limitations are, perhaps, more eye opening. Many empirical studies of consumer demand have associated a share format to the inconsistent assumption that disturbance terms are multivariate normally distributed. The consequence is that some predicted shares may be negative, especially if the observed share is close to zero (Aitchison, 1982; Fry *et al*., 1996). Another limitation is the impossibility of testing a crucial null hypothesis such as the adding-up condition which is automatically satisfied in an expenditure-share format and induces the singularity of the error covariance matrix. In a share format, adding-up, symmetry and homogeneity are hypotheses that cannot be tested independently. Testing the adding-up condition is important because, often, the number of sample commodities is much smaller than the number of goods that compose a consumer's basket. This hypothesis constitutes the paper's main focus.

The advantages of a quantity format can be listed as the admission of a multivariate normal structure of the disturbance terms, the possibility of testing the adding-up condition, the zero-degree homogeneity assumption and the symmetry and negative semidefiniteness of the Slutsky matrix as separate null hypotheses. The disadvantages are minimal and deal, possibly, with the necessity of requiring larger samples than in the case of a share format. This event may occur in very small samples.

Given the gamut of issues associated with the estimation and testing of consumer demand systems, we will narrow the discussion to specifications of share systems as commonly appeared in the literature. The pioneering paper by Sir Richard Stone (1954, p. 512) presents a linear expenditure system (LES) of demand functions stated in expenditure format, where the dependent variable represents the expenditure on a given good. This specification is equivalent to a share format where the share is defined with respect to total expenditure. Stone's LES empirical model includes all goods and services grouped in six categories of commodities for the years 1920 to 1938 in the United Kingdom. For the first time, the theoretical requirements of adding-up, zero-degree homogeneity of demand functions and symmetry of the Slutsky matrix appear as restrictions in the empirical literature. A. P. Barten (1964), who presented a linear demand system stated directly in share format, attempted to include all commodities in the consumer expenditure household survey kept in The Netherlands between 1921 and 1958. There followed other important papers by Barten (1968, 1969) in share format and by Pollak and Wales (1969) in expenditure format. Hence, the tradition of estimating demand systems in expenditure-share format has a distinguished lineage.

In his influential paper that summarizes the empirical literature on consumer demand, Barten (1977, page 23) wrote: "The approach is essentially an empirical one, in the sense that one aims at the formulation of a system to be estimated using actual data. In view of the data limitations, one makes use of restrictions which, in part, are of a theoretical nature." We interpret Barten's words to mean that the data generating process (DGP) ought to assume center stage in an econometric specification of models that wishes to represent the final decisions of consumer behavior. In econometrics, a DGP must be guided by economic theory but must also be adapted to describe the peculiarities of data collection, as Barten implicitly suggests.

In the case of consumer behavior, utility theory develops the process of deriving systems of demand functions in the format of quantity levels of various commodities as a function of their prices and income. Let q be an N-vector of quantities of N commodities and services that represent all the goods' choices available to a consumer. Let p be an N-vector of prices of those goods. Finally, let m be the exogenous income available to consumer for making her N decisions. Then, utility theory derives a system of $(N+1)$ relations that are intepreted as N Marshallian demand functions and a budget constraint

$$q = q(m, p) \tag{1}$$

$$p'q = m \; . \tag{2}$$

The $(N+1)$ system has N unknown quantities, q, and, therefore, one of the N relations in (1) is redundant and can be omitted in the solution of the remaining $(N-1)$ quantities. The quantity of the Nth good can be recovered from the budget constraint after replacing the $(N-1)$ quantities obtained from the solution of the $(N-1)$ relations.

In many cases, however, the DGP of consumer demand information, in any given sample, may not satisfy all the conditions stated above. Many empirical studies that estimate systems of demand functions exhibit a number of commodities, $n < N$, that is much smaller than the number of all possible goods available for consumers' decisions over a given time interval. In this case, the sample demand system is incomplete: It does not satisfy the adding-up condition although the demand functions are still (theoretically) homogeneous of zero-degree in prices and income.

It is well known that, to justify the adoption of the features associated with the general theoretical scaffolding also in the case of a very small number of commodities (or commodity aggregates), the hypotheses of separability and multistage budgeting were developed. Accordingly, consumption decisions would occur in at least two stages. In the first stage, consumer would allocate income among a number of commodity subsets. In the second stage, consumer would proceed to maximize utility only with respect to the commodities belonging to one of those subsets subject to the previously determined portion of income for that category of goods. All this is well from a theoretical standpoint. In general, however, these hypotheses remain untested and untestable, given the available sample information. Put another way, the portion of income that, according to a two-stage approach of consumer decisions, would be allocated to a specific commodity subset in the first stage is never known and measurable, thus invalidating the assumption that would require this level of income to be an exogenous piece of information.

As a consequence, in empirical studies, the budget constraint (2) may never bind. Furthermore, information on total exogenous income is never collected. What Barten calls total expenditure, m, is simply an accounting definition analogous to (2) but generated as the sum of sample prices times quantities over the available n commodities. Often, therefore, for econometric purposes, there are only n independent equations similar to (1) while the analogous equation (2) is not a constraint but is simply an accounting relation with no sample information of its own that is independent of prices and quantities. Many empirical studies of demand published to date, however, have taken for valid both relations (1) and (2), regardless of the subset of commodities dealt with in the sample and without performing a statistical test of the adding-up condition. This test appears to be crucial for assessing the theoretical scaffolding leading to a share format: If constraint (2) is part of the untestable hypothesis that the sample commodities constitute a proper subset of goods within a two-stage budgeting process, the test of adding-up condition is an indicator of whether that hypothesis may be supported by the sample data.

Referring to a stochastic specification of a demand system described by the theoretical scaffolding (1) and (2), the fundamental, empirical consequence of the assumptions and conclusions that are valid for the entire consumer's basket is stated by Barten (1977, p. 26) as: "However, (2) implies a linear dependence of the joint distribution of the disturbances if m and p are exogenous. The theoretical covariance matrix is, therefore, singular. This problem is usually solved by deleting one equation from the system."

This proposition was originally put forward in the late sixties (Barten, 1968, 1969; Pollak and Wales, 1969) and, since then, almost all the empirical studies of demand that appeared in the literature have adopted it, regardless of the number of commodities involved. Furthermore, the great majority of studies have gone another step and have specified demand systems in the format of expenditure shares. Deaton and Muellbauer (1980), with their Almost Ideal Demand System (AIDS), have provided a remarkable impetus for the use of an expenditure-share format in empirical studies of demand. To repeat the question that forms the objective of this paper: Should all empirical demand studies – even those that that are based upon a few commodities – adopt a share format?

Thus, in this cursory survey of empirical demand issues, we have identified two main topics of interest. The first topic deals with the question whether the DGP of sample information of consumer behavior – as typically observed – statistically supports the application of the more general approach embedded in equations (1) and (2), regardless of the size of the subset of commodities constituting the sample data. The second topic discusses the consequences of estimating demand systems in expenditure-share format rather than in a quantity format. In particular, given the absence of empirical information about two-stage budgeting and separability that characterizes many empirical demand studies, it is of interest to know whether the adding-up condition holds for the sample at hand. As elaborated in more detail further on, this condition is crucial for concluding that the error covariance matrix is singular and, as a consequence, for admitting the deletion of an equation in the estimation of demand parameters without loss of information. The adding-up condition, however, cannot be tested using an expenditure-share format of the demand system. This test must be perfomed using a quantity format.

The paper is organized in several sections. Section 2 presents a general discussion of estimating models (not necessarily models of consumer behavior) in a share format. Section 3 discusses the issue of estimating systems of demand functions in a quantity format when the number of commodities in the sample is less than the number of commodities in the consumer basket, that is, $n < N$. Section 4 lays out the stochastic quantity model of demand functions based upon the AIDS specification of Deaton and Muellbauer (1980) and a series of null hypotheses that are relevant for validating (or refuting) the model specification. Section 5 describes a sample of cross-section data used in the empirical analysis and presents the empirical results. Conclusions follow.

2. Models in Share Format

Any linear statistical model that is specified in share format, with an intercept in each equation and the same explanatory variables appearing in every equation, exhibits a unique property: the sum over equations of the least-squares (LS) estimated residuals is equal to zero in each sample observation. Therefore, the estimated error covariance matrix is singular. Furthermore, the sum over intercepts of the various equations is equal to 1 and the sum over rows of the coefficient matrix associated with explanatory variables is equal to zero without any *a priori* condition on parameters. Hence, the adding-up property of shares holds automatically on the left and on the right of the equality sign. This result appears in papers by Worswick and Champernowne (1954-1955), Barten (1969), Berndt and Savin (1975) and Edgerton *et al*. (1996, ch. 11). We offer an alternative derivation in the Appendix. Surprisingly, however, many demand studies[1] that specify a share format declare that the adding-up restrictions must be imposed on the model's parameters. This oversight may have consequences for testing hypotheses.

Let $t = 1,...,T$ indicate sample observations; $k = 1,..,K$ the number of equations; $j = 1,...,J$ the number of explanatory variables; w_{kt} the share of the *k*th equation in the *t*th observation; p_{jt} the *j*th explanatory variable in the *t*th observation; b_k the intercept in the *k*th equation; a_{jk} the *j*th parameter in the *k*th equation; u_{kt} the disturbance term of the *k*th equation in the *t*th observation with expectation $E(u_{kt}) = 0$ and constant $(K \times K)$ contemporaneous covariance matrix Σ_u. All explanatory variables appear in each equation. Then, a share model without theory is stated as

$$w_{kt} = b_k + \sum_{j=1}^{J} a_{jk} p_{jt} + u_{kt} . \tag{3}$$

Summing over equations

$$
\begin{aligned}
\sum_{k=1}^{K} w_{kt} &= \sum_{k=1}^{K} b_k + \sum_{k=1}^{K}\sum_{j=1}^{J} a_{jk} p_{jt} + \sum_{k=1}^{K} u_{kt} \\
1 &= \sum_{k=1}^{K} b_k + \sum_{k=1}^{K}\sum_{j=1}^{J} a_{jk} p_{jt} + \sum_{k=1}^{K} u_{kt} .
\end{aligned}
\tag{4}
$$

For the class of least-squares estimators that minimize the sum of squared residuals, it can be shown (Appendix) that – in each observation – the sum over equations of the estimated residuals in (4) is equal to zero and the adding-up property is fulfilled automatically without imposing *a priori* any additional constraints on the parameters of the share model specification (3). In other words,

$$
\begin{aligned}
\sum_{k=1}^{K} w_{k} &= \sum_{k=1}^{K} \hat{b}_k + \sum_{k=1}^{K}\sum_{J=1}^{J} \hat{a}_{jk} p_{jt} + \sum_{k=1}^{K} \hat{u}_{kt} \\
1 &= \;1 \quad + \qquad 0 \qquad + \quad 0.
\end{aligned}
\tag{5}
$$

The estimated and contemporaneous residuals $\hat{u}_{kt}$ form a linear combination in each observation and the estimated error covariance matrix is singular. Therefore, any estimator that requires the inversion of the covariance matrix $\hat{\Sigma}_u$ is infeasible. Notice that

$$
\sum_{k=1}^{K} \hat{b}_k = 1 \qquad \text{and} \qquad \sum_{k=1}^{K} \hat{a}_{jk} = 0, \quad j = 1,...,J \tag{6}
$$

without the necessity to impose these conditions as *a priori* restrictions. Hence, an equation can be deleted from (3) and the estimates of the corresponding parameters can be recovered from relations (6).

The relationship between this discussion of a general share system such as (3) and an expenditure-share system of demand functions, as usually stated in the literature, is straightforward. Many demand studies appeared in print and specified in expenditure-share format – although they deal with a number of commodities $n < N$ – have all explicitly assumed and imposed adding-up conditions by way of parameter restrictions analogous to (6). But since the adding-up condition holds by necessity without the need to impose it *a priori*, this suggests that the share specification of any econometric model (and, equivalently, the expenditure specification of it) is like a straight jacket: once worn, it forces the error covariance matrix to be singular and the adding-up condition to hold whether or not the DGP warrants it. An important corollary follows: the null hypothesis that the adding-up condition holds cannot be tested under a share (expenditure) format of demand systems. In the absence of any sample information regarding two-stage budgeting, the test of the null hypothesis that the adding-up condition holds corresponds to an indirect test of the assumption that the sample commodities constitutes a proper subset of goods in a two-stage budgeting process of consumer behavior. To test this null hypothesis, however, only a quantity format specification of a demand system is available.

To exemplify more directly that the above reasoning applies also to demand systems, we state the AIDS model of Deaton and Muellbauer (1980) in share format

$$w_{kt} = \alpha_k + \sum_{i-1}^{n} \gamma_{ki} \log p_{it} + \beta_k \log\left[\frac{x_t}{P_t}\right] + \mathrm{u}_{kt} \tag{7}$$

where $k = 1,...,n,\ i = 1,...,n$ and $t = 1,...,T$. There are $n < N$ commodities with q_{kt} and p_{kt} representing quantities and prices of the *t*th sample observation while total expenditure is $x_t = \sum_{k=1}^{n} p_{kt} q_{kt}$ with shares computed as $w_{kt} = p_{kt} q_{kt} / x_t$. Furthermore, the deflating price index is defined as

$$\log P_t = \alpha_0 + \sum_{i=1}^{n} \alpha_i \log p_{it} + \frac{1}{2} \sum_{i=1}^{n} \sum_{k=1}^{n} \gamma_{ik} \log p_{it} \log p_{kt} \tag{8}$$

although Deaton and Muellbauer suggested – and many empirical studies adopted their suggestion – that a Stone index could often suffice:

$$\log P_t^* = \sum_{i=1}^{n} w_{it} \log p_{it}. \tag{9}$$

Furthermore, they specify and impose parameter restrictions that include adding-up requirements, zero-degree homogeneity in prices and income of demand functions and symmetry of the Slutsky matrix

$$\sum_{k=1}^{K} \alpha_k = 1, \quad \sum_{k=1}^{K} \gamma_{ki} = 0, \quad \sum_{k=1}^{K} \beta_k = 0 \qquad \text{adding-up} \tag{10}$$

$$\sum_{i=1}^{K} \gamma_{ki} = 0 \qquad \text{zero-degree homogeneity} \tag{11}$$

$$\gamma_{ki} = \gamma_{ik} \qquad \text{Slutsky symmetry} \tag{12}$$

and write (1980, p. 314): "Provided (10), (11), and (12) hold, equation (7) represents a system of demand functions which add up to total expenditure ($\sum_k w_k = 1$), are homogeneous of degree zero in prices and total expenditure taken together, and which satisfy Slutsky symmetry." But, as argued above, restrictions (10) are automatically satisfied in a share system regardless of either theory or other assumptions. They are satisfied automatically also when conditions (11) and (12) are imposed using either specification of the price index deflator. Hence, there is no need to state them as if they "ought to be imposed" for estimating a share model which represents a demand system.

Thus, the estimation of equations (7) and (8) [or (9)] together with side conditions (11) and (12) represents a special case of estimating the share system (3). Barten (1969, p. 16) stated: "… it is possible to delete one equation from the system without losing any information."[2] After the knowledge acquired from the above discussion, this statement should be qualified to read: “When a share format is warranted, it is possible to delete one equation from the system without losing any information.”

With respect to parameter “restrictions” (10) a crucial remark is in order. They imply that the general theoretical conclusions of consumer theory, which are valid for the full basket of N commodities, have been adopted also for the case when the number of sample goods is $n < N$. Second, the symmetry of the gamma parameters and the zero-degree omogeneity restriction imply the adding-up condition. Furthermore, by itself, the adding-up hypothesis cannot be tested in an expenditure-share demand system.

Suppose that the adding up condition does not hold (tested in a quantity format model). This means that the number of sample commodities is different from the number of goods constituting a proper subset, according to a two-stage budgeting criterion. Therefore, the zero-degree homogeneity condition cannot have the form stated in (11). The n equations may still be homogeneous of degree zero in prices and income but the missing prices – corresponding to the missing commodities – prevent the use of (11). Zero-degree homogeneity of demand functions depend on relative prices. Therefore, even if the adding-up condition does not hold, the sample commodities may belong to a set of functions where the zero-degree homogeneity condition holds. A test of this condition in the sample at hand requires a modification of the functional form of the demand functions, as explained below.

3. Models in Quantity Format

To further motivate the research objective of this paper, it is of interest to analyze in some detail Pollak and Wales’ 1969 paper from a viewpoint that, in many cases, (*i*) the number of commodities of a consumer sample is rather limited and does not exhaust all the consumer choices; (*ii*) it is unknown and untestable whether consumers regard those commodities as a proper subset in a two-stage budgeting procedure; (*iii*) it is unknown and untestable whether the budget constraint (2) holds for the given sample information.

The relevance of Pollak and Wales 1969 paper stems from their demonstration that – under a specific assumption – the error covariance matrix of a LES demand system specified in quantity format is singular. When this scenario holds, one can restate the demand system in either an expenditure or a share format by dropping one equation without loss of information. But, as we have repeatedly suggested, it is difficult to know *a priori* whether, given a sample of only a few commodities, the budget constraint holds under a two-stage budgeting: the ideal empirical process would be to test this hypothesis. On the other hand, many data samples do not allow a direct test. The test of the adding-up condition, then, is an alternative test of the null hypothesis that the given sample commodities constitutes a proper subset in a two-stage budgeting framework.

Pollak and Wales (1969, p. 618) – who dealt with a sample of four commodity groups – ended up estimating an LES demand system in expenditure format that, we know, imposes a singular error covariance matrix. They justified this estimation procedure by demonstrating that also the error covariance matrix of the same demand system – specified in quantity format – is singular. In Pollak and Wales' notation, the LES model generates demand functions of the following matrix specification (Pollak and Wales, 1969, p. 619)

$$x = (I - \gamma p')b + \gamma\mu \qquad (13)$$

where x is an $(n \times 1)$ vector of quantities, p is an $(n \times 1)$ vector of prices, μ is a scalar equal to total expenditure, b is an $(n \times 1)$ vector of parameters of the LES utility function to be estimated, and γ is an $(n \times 1)$ vector with elements a_i / p_i where the a_i are parameters of the LES utility function to be estimated and normalized to $\sum_i a_i = 1$. Pollak and Wales write (1969, p. 619): "The stochastic specification which appears to us most appropriate … assumes that the disturbances are associated with the b's." Hence, they write: "

$$\begin{aligned} x &= (I - \gamma p')(b + u) + \gamma\mu \\ &= (I - \gamma p')b + \gamma\mu + (I - \gamma p')u \\ &= (I - \gamma p')b + \gamma\mu + v \end{aligned} \tag{14}$$

where the vector u is assumed to be multivariate normal with mean 0 and full-rank diagonal covariance matrix D; the error $v = (I - \gamma p')u = Mu$ is a linear transformation of u with covariance matrix given by $\Omega = MDM'$. Since the transformation M is singular (due to the restrictions on the a_i parameters), the covariance matrix Ω is singular." It appears that without the assumption that "the disturbances are associated with the b's" – in a quantity format specification – it would not have been possible to show that Ω is singular.

Notice also that, in their demonstration, Pollak and Wales did not use the adding-up condition on expenditures, but relied on the normalization of the a_i parameters, to establish the singularity of the M transformation. This suggests that the LES system of demand functions in its quantity specification does not, by itself, imply that the corresponding error covariance matrix is singular. To show its singularity it is necessary to transform the system into an expenditure specification which makes use of the adding-up condition represented by the budget constraint. This transformation, however, generates the following result: With sample total expenditure defined as $\sum_k p_k q_k = m$

and the stochastic demand function as $q_k = q_k(m,p) + v_k = E(q_k) + v_k$, one would expect that the disturbance terms of commodity quantities be trasmitted to total expenditure. And yet, as Pollak and Wales elaborate (1969, p. 615), the transformation of the demand system into an expenditure system will have that

$$\begin{aligned} \sum_{k=1}^{K} p_k q_k &= m \\ \sum_{k=1}^{K} p_k [E(q_k) + v_k] &= m \\ \sum_{k=1}^{K} p_k E(q_k) + \sum_{k=1}^{K} p_k v_k &= m \end{aligned} \tag{15}$$

and since an expenditure format is equivalent to a share format with a singular covariance matrix, it must be that $\sum_k p_k v_k = 0$, with the result that the disturbance terms of the commodity quantities are no longer transmitted to total expenditure. Therefore, $m = \sum_k p_k q_k = \sum_k p_k E(q_k)$, with $q_k \neq E(q_k)$. Of course, a scalar may correspond to an infinite number of linear combinations. In a quantity specification, however, the estimated residuals do not satisfy, in general, a linear combination of prices (unless we know that the sample commodities exhaust consumer's basket goods).

The quantity format of the demand model, therefore, is the proper specification for testing the null hypothesis *H*0 that the adding-up condition holds, because the corresponding covariance matrix is not necessarily singular and the likelihood function is in general well defined also under *H*0.

If *H*0 is not rejected, the sample information supports the hypothesis that the commodities involved constitute a proper subset of goods that is consistent with a two-stage budgeting of consumer behavior. In this case it is possible to respecify the model in share format and to delete one equation without loss of information, although – at this stage – a legitimate question is: Why repeating the estimation of the demand system? If $H0$ is rejected, the demand model should ultimately be estimated in quantity format without imposing the adding-up restrictions. A test for zero-degree homogeneity requires a modification of the functional form. A share format, in this case, would imply loss of information.

4. AIDS Quantity Form

Under the assumptions of an AIDS expenditure function, consumer utility theory generates a system of demand functions that assumes the following quantity format in a stochastic representation

$$q_{it} = \alpha_i \frac{x_t}{p_{it}} + \sum_{j=1}^{n} \gamma_{ij} \log p_{jt} \frac{x_t}{p_{it}} + \beta_i \log\left[\frac{x_t}{P_t}\right] \frac{x_t}{p_{it}} + v_{it} g_i(x_t, \mathbf{p}_t) \tag{16}$$

where $i = 1,...,n,\ j = 1,...,n$ and v_{it} is a disturbance term for the *i*th commodity in the *t*th observation with expectation $E(v_{it}) = 0$ and covariance matrix Σ_v. According to Brown and Walker (1989) the disturbance terms of commodities involved in the individual consumer's decisions may depend on prices and total expenditure. To represent this assumption about heteroskedasticity the function $g_i(x_t, p_t)$ multiplies the disturbance term with the objective of rendering v_{it} homoskedastic.

Model (16) can now be used to test a series of null hypotheses based upon restrictions (10), (11) and (12). The tests have the structure of a likelihood ratio which is distributed as a chi square with degrees of freedom equal to the number of restrictions. The first step is to set up an overall null hypothesis that deals with all the parameter restrictions in (10), (11) and (12). Notice, however, that (10) and (12) imply (11). Therefore, the formal statement to test is

Overall Null Hypothesis: $H0_o$:(10) and (12), $H1_o$: not $H0_o$.

If $H0_O$ is not rejected, the ideal next step would be to test the null hypothesis that the error covariance matrix Σ_v is singular. This test is rather complex and, to date, no clear procedure has appeared in the literature. If $H0_O$ is rejected, it is of interest to investigate in more detail the source of rejection. Toward this objective, a sequence of hypotheses are formulated.

Adding-up: $H0_A$:(10); $H1_A$: not $H0_A$

If $H0_A$ is not rejected we repeat here that the ideal next step would be to test the null hypothesis that the error covariance matrix Σ_v is singular. In a quantity format demand system, the numerical estimate of the error covariance matrix is likely to produce a non-singular matrix even when $H0_A$ is not rejected. In this case, a proper procedure would call for the test that the smallest eigenvalue is equal to zero. To date, this test is not in the toolkit of econometricians. Assuming, therefore, that the covariance matrix is singular under the non rejection of $H0_A$, it is admissible to re-estimate the demand system (16) in share format because no loss of information will occur. But, at this point, there is no longer reason to do that. If $H0_A$ is rejected, an expenditure-share format of the demand model is definitely unwarranted. In this case, the wrongful use of a share format and the drop of an equation would correspond to a loss of information because the error covariance matrix is not singular.

Slutsky Symmetry and Zero-Degree Homogeneity

The fundamental conclusions of the traditional consumer theory are captured by the symmetry and the negative semi-definiteness of the Slutsky matrix and by the absence of money illusion. The latter property is commonly associated with a zero-degree homogeneity of the demand functions with respect to all prices and income. The two conditions go together and they ought to be tested in a coupled form. However, if the adding-up condition and symmetry hold, the homogeneity hypothesis cannot be tested even in a quantity format specification. It will have to be considered a maintained hypothesis since adding-up and symmetry imply homogeneity. In this case, the null hypothesis is articulated as $H0_{SH}$:(11) and (12); $H1_{SH}$: not $H0_{SH}$

A principal submatrix of a symmetric negative semidefinite matrix is itself symmetric and negative semidefinite. Hence, the symmetry condition (12) may hold even though the adding-up condition may not. The Slutsky matrix S is symmetric negative semidefinite and, under AIDS, takes the form (Moschini, 1998)

$$S_{ki} = \frac{x}{p_k p_i}\left[\gamma_{ki} + w_k w_i - \delta_{ki} w_k + \beta_k \beta_i \log\left(\frac{x}{P}\right)\right]$$

where δ_{ki} is the Kronecker delta ($\delta_{ki} = 1$ for $k = i$ and $\delta_{ki} = 0$ for $k \neq i$). Hence, the symmetry of the gamma parameters is required for guaranteeing the symmetry of the Slutsky matrix.

Rejection of $H0_{SH}$ constitutes a severe blow to the theory that generated the AIDS demand system because the symmetry (and negative semidefiniteness) of the Slutsky matrix and the zero-degree homogeneity of demand functions constitute the fundamental signature of consumer theory.

Notice that the rejection of $H0_A$ implies that the adding-up condition does not hold. As a consequence, the zero-degree homogeneity hypothesis cannot be evaluated in terms of restrictions (11) because the sample commodities do not constitute a proper subset of goods. As the homogeneity property is based upon relative prices, demand functions may still be homogeneous of zero degree in prices and income even though the adding-up condition does not hold. Therefore, model (16) is respecified in terms of relative prices by dividing each nominal price by the price index P.

$$q_{it} = \alpha_i \frac{x_t}{p_{it}} + \sum_{j=1}^{n} \gamma_{ij} \log\left[\frac{p_{jt}}{P_t}\right]\frac{x_t}{p_{it}} + \beta_i \log\left[\frac{x_t}{P_t}\right]\frac{x_t}{p_{it}} + \nu_{it} g_i(x_t, \mathbf{p}_t)\,. \qquad (17)$$

Slutsky Symmetry & Zero-Degree Homogeneity when the adding-up condition fails:

$H0_{SH|notA}$: (12); $H1_{SH|notA}$: not $H0_{SH|notA}$

The test of this hypothesis is based upon the estimation of relation (17). Hence, the test of Slutsky symmetry is conditional on the zero-homogeneity of the AIDS demand functions. If $H0_{SH|notA}$ is not rejected, the sample data support the conclusions of the traditional consumer theory. If $H0_{SH|notA}$ is rejected, the AIDS theory is refuted and other suitable specifications of the demand system may be investigated.

5. Data and Results

The estimation and hypothesis testing strategy outlined in previous sections is applied to a cross-section sample of 119 consumers who purchased four food commodities: bread and cereals, meat, beverages, other foods. Information on quantities and expenditures on the four goods is available for each sample unit.[3] Hence, it is possible to compute the corresponding commodity prices (unit values) for each consumer.

We remind readers that the principal objective of this paper consists in testing the appropriateness of either a share or a quantity format in the estimation of demand systems. The estimated AIDS model is expressed by relation (16), with a switching to relation (17) in case the adding-up hypothesis is refuted. The parameter α_0 in the price index (8) was fixed at the logarithm of the minimum total expenditure value in the sample minus a small constant, as suggested by Banks *et al*. (1997, p. 534). The heteroskedasticity of errors is specified by selecting $g_i(p_t, x_t)$ = total consumer expenditure. All tests are in the structure of a likelihood ratio.

The first null hypothesis concerns the validity of all restrictions taken together. This means that the null hypothesis is defined as the set of restrictions specified in relations (10) and (12) – that imply (11). This hypothesis is rejected. The chi squared variable, with 12 degrees of freedom, soundly rejects the null hypothesis also at 1 percent confidence level (see Table 1).

The next step deals with testing the adding-up restrictions as represented by relation (10). Also this hypothesis is rejected with high confidence. With 6 degrees of freedom, the corresponding chi squared achieves a value of 45.76 > 16.81, the critical value at 1 percent confidence level. A share format is not warranted for this data sample.

(Table 1 here)

To complete the test of consumer theory there remains to verify whether the matrix of the gamma parameters is symmetric. As homogeneity is a built-in property of the traditional consumer theory, we switch to relation (17) which incorporates the absence of money illusion by means of relative prices. Hence, the test of Slutsky symmetry is conditional upon the maintained hypothesis that the model exhibits zero-degree homogeneity. The symmetry of the Slutsky matrix is refuted (see Table 1). As a consequence, the AIDS specification is not a suitable framework for representing the behavior of this consumers' sample.

The parameter estimates for the symmetry-restricted specification under relative prices (model (17)) are reported in Table 2. The alpha coefficients add up to 1.075 and the sum of the beta parameters is equal to -0.008. The gamma matrix is symmetric but its row and columns do not add up to a zero value.

(Table 2 here)

6. Conclusion

This paper's motivation sprang from the question of whether a share format of demand systems is warranted even in cases when the data sample deals with a rather small number of consumer goods. The adding-up condition was identified as a crucial restriction that may not be attained when demand systems are incomplete. In such cases, the error covariance matrix of the empirical model (specified in quantity format) is not singular and a share format is unwarranted because dropping an equation – as customarily done in the estimation of share specifications – corresponds to losing sample information.

The estimation of quantity formats does not involve any additional difficulties over those ones encountered in the estimation of share formats. Quantity formats, furthermore, allow for testing all the relevant hypotheses of consumer theory, including the adding-up restriction – an hypothesis that is precluded by share formats.

The illustration of the research strategy discussed in the paper dealt with a cross-section sample of 119 consumers and four food commodities. All the null hypotheses were refuted, as indicated by the results of Table 1. This means two things. First, having chosen a traditional AIDS specification, the sample data failed to support the adding-up condition. A share format, therefore, should not be used with this specification. Second, also the Slutsky symmetry (conditional upon the maintained hypothesis of zero-degree homogeneity of the demand functions) was refuted, indicating that the sample data do not support the AIDS specification. The next step ought to be the selection of alternative econometric specifications, something left for future research. We stop at this stage because we wish to leave the reader with an emphasis on the main point of the paper: share versus quantity formats.

Appendix

We wish to show that, for each observation, the sum over equations of least-squares estimated residuals of any seemingly unrelated equation system in share format (with intercepts and the same explanatory variables entering every equation) is equal to zero. As a consequence, the associated error covariance matrix is singular.

Let $k = 1,...,K$ be the number of equations in share format; $t = 1,...,T$ be the number of observations in each equation; $W = [w_1, w_2, ..., w_K]$ be a matrix of K $(T \times 1)$ vectors of shares; $P = [p_1, p_2, ..., p_J]$ be a matrix of J $(T \times 1)$ vectors of explanatory variables that enter each equation; $U = [u_1, u_2, ..., u_K]$ be a matrix of K $(T \times 1)$ vectors of disturbance terms; s_T be an $(T \times 1)$ vector of unitary values; s_K be a $(K \times 1)$ vector of unitary values; b be a $(K \times 1)$ vector of intercepts; A be a $(J \times K)$ matrix of unknown parameters.

Given these stipulations, a share system can be specified in matrix form as

$$W = s_T b' + PA + U \tag{A1}$$

or, more compactly,

$$W = \begin{bmatrix} P & s_T \end{bmatrix} \begin{bmatrix} A \\ b' \end{bmatrix} + U \ . \tag{A2}$$

Let $X = \begin{bmatrix} P & s_T \end{bmatrix}$ and $B' = \begin{bmatrix} A' & b \end{bmatrix}$. Then, system (A2) can be represented as

$$W = XB + U \ . \tag{A3}$$

Assuming that the inverse matrix $(X'X)^{-1}$ exists, the least-squares estimate of matrix B is given as $\hat{B} = (X'X)^{-1} X'W$, with residuals

$$U = \left[I - X(X'X)^{-1} X' \right] W \ . \tag{A4}$$

Using the sum vector s_K on matrices U and W :

$$\begin{aligned} Us_K &= [I - X(X'X)^{-1} X'] W s_K \\ Us_K &= [I - X(X'X)^{-1} X'] s_T \\ Us_K &= s_T - X(X'X)^{-1} X' s_T \end{aligned} \tag{A5}$$

Recall that

$$X' s_T = \begin{bmatrix} P' \\ s_T' \end{bmatrix} s_N = \begin{bmatrix} P' s_T \\ s_T' s_T \end{bmatrix}$$

$$(X'X)^{-1} = \left\{ \begin{bmatrix} P' \\ s_T' \end{bmatrix} \begin{bmatrix} P & s_T \end{bmatrix} \right\}^{-1} = \left\{ \begin{matrix} P'P & P's_T \\ s_T'P & s_T's_T \end{matrix} \right\}^{-1} .$$

From Hadley (1962, p. 36), the inverse of a partitioned matrix, in general form, can be computed as

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} I_1 & 0 \\ 0 & I_2 \end{bmatrix}$$

where $A=(\alpha-\beta\delta^{-1}\gamma)^{-1}$, $B=-A\beta\delta^{-1}$, $C=-\delta^{-1}\gamma A$ and $D=\delta^{-1}-\delta^{-1}\gamma B$.

In the share case, however, $\gamma=\beta'$. Hence, the inverse matrix corresponding to $(X'X)^{-1}$ is given by

$$\begin{bmatrix} A & B \\ B' & D \end{bmatrix}=\begin{bmatrix} (\alpha-\beta\delta^{-1}\beta')^{-1} & -A\beta\delta^{-1} \\ -\delta^{-1}\beta' A & \delta^{-1}-\delta^{-1}\beta' B \end{bmatrix} \tag{A6}$$

With these preliminaries,

$$(X'X)=\begin{Bmatrix} P'P & P's_T \\ s_T'P & s_T's_T \end{Bmatrix}=\begin{bmatrix} \alpha & \beta \\ \beta' & \delta \end{bmatrix} \tag{A7}$$

with an inverse matrix whose components are

$$\begin{Bmatrix} P'P & P's_T \\ s_T'P & s_T's_T \end{Bmatrix}^{-1}=\begin{bmatrix} \left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1} & -\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}\dfrac{P's_T}{s_T's_T} \\ -\dfrac{s_T'P}{s_T's_T}\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1} & \left(\dfrac{1}{s_T's_T}+\dfrac{s_T'P}{s_T's_T}\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}\dfrac{P's_T}{s_T's_T}\right) \end{bmatrix}$$

Therefore,

$$(X'X)^{-1}X's_T=\begin{Bmatrix} P'P & P's_T \\ s_T'P & s_T's_T \end{Bmatrix}^{-1}\begin{bmatrix} P's_T \\ s_T's_T \end{bmatrix}$$

$$=\begin{bmatrix} \left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}P's_T & -\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}P's_T \\ -\dfrac{s_T'P}{s_T's_T}\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}P's_T+ & 1+\dfrac{s_T'P}{s_T's_T}\left\{P'P-\dfrac{P's_Ts_T'P}{s_T's_T}\right\}^{-1}P's_T \end{bmatrix}=\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Finally, recovering the last equation of (A5)

$$Us_K = s_T - X(X'X)^{-1}Xs_T$$
$$= s_T - X\begin{bmatrix} 0 \\ 1 \end{bmatrix} = s_T - \begin{bmatrix} P & s_T \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \qquad \text{Q.E.D.} \qquad \text{(A8)}$$

Furthermore, the sum over equations of the intercepts coefficients is equal to 1 while the sum over equations of the coefficients of the A matrix is equal to zero for each column of A. These results are shown below. Recall that $\hat{B} = (X'X)^{-1}X'W$. Then, summing over the index $k = 1,...,K$

$$\hat{B}s_K = (X'X)^{-1}X'Ws_K = (X'X)^{-1}X's_T$$
$$\begin{bmatrix} \hat{A}s_K \\ \hat{b}'s_K \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \qquad \text{(A9)}$$

References

Aitchison, J., (1982). The Statistical Analysis of Compositional Data. *Journal of the Royal Statistical Society B* 44: 139-177.

Alston, J. M., Chalfant, J. A. and Piggott, N. E. (2001). Incorporating Demand Shifters in the Almost Ideal Demand System. *Economics Letters* 70: 73-78.

Banks, J., Blundell, R. and Lewbel, A. (1997). Quadratic Engel Curves and Consumer Demand. *Review of Economics and Statistics* 79: 527-539.

Barnett, W. A. and Serletis, A. (2008). Consumer Preferences and Demand Systems. *Journal of Econometrics* 147: 210-224.

Barten, A. P. (1964). Consumer Demand Functions under Conditions of Almost Additive Preferences. *Econometrica* 32: 1-38.

Barten, A. P. (1968). Estimating Demand Functions. *Econometrica* 36: 213-251.

Barten, A. P. (1969). Maximum Likelihood Estimation of a Complete System of Demand Equations. *European Economic Revie*w 1: 7-73.

Barten, A. P. (1977). The System of Consumer Demand Functions Approach: A Review. *Econometrica* 45: 23-50.

Berndt, E. R. and Savin, N. E. (1975). Estimation and Hypothesis Testing in Singular Equation Systems with Autoregressive Disturbances. *Econometrica* 43: 937-958.

Brown, B. W. and Walker, M. B. (1989). The Random Utility Hypothesis and Inference in Demand Systems. *Econometrica* 57: 815-829.

Cranfield, J. A. L., Eales, J. S., Hertel, T. W. and Preckel, P. V. (2003). Model Selection when Estimating and Predicting Consumer Demands Using International, Cross Section Data. *Empirical Economics* 28: 353-364.

Deaton, A. and Muellbauer, J. (1980). An Almost Ideal Demand System. *The American Economic Review* 70: 312-326.

Edgerton, D., Assarsson, B., Hummelmose, A., Laurila, I. P., Rickertsen, K. and Vale, P. H. (1996). *The Econometrics of Demand Systems With Applications to Food Demand in the Nordic Countries* (Dordrecht, NL: Kluwer Academic Publishers).

Fisher, D., Fleissig, A. R. and Serletis, A. (2001). An Empirical Comparison of Flexible Demand System Functional Forms. *Journal of Applied Econometrics* 16: 59-80.

Fry, J. M., Fry, T. R. L. and McLaren, K. R. (1996). The Stochastic Specification of Demand Share Equations: Restricting Budget Shares to the Unit Simplex. *Journal of Econometrics* 73: 377-385.

Hadley, G. (1962). *Linear Programming* (Reading, MA: Addison-Wesley).

Moschini, G. (1998). The Semiflexible Almost Ideal Demand System. *European Economic Review* 42: 349-364.

Pollak, R. A. and Wales, T. J. (1969). Estimation of the Linear Expenditure System. *Econometrica* 37: 611-628.

Stone, R. (1954). Linear Expenditure Systems and Demand Analysis: An Application to the Pattern of British Demand. *The Economic Journal* 64: 511-527.

Worswick, G. D. N. and Champernowne, D. G. (1954-1955). A Note on the Adding-up Criterion. *The Review of Economic Studies* 22: 57-60.

Footnotes

[1] For example, Berndt and Savin (1975, p. 938) write: "It is assumed that *y* satisfies the adding-up conditions…"; Moschini (1998, p. 351) writes: "… adding-up … hold(s) if …"; Alston, Chalfant and Piggott (2001, p. 74) write: "To satisfy … adding-up …the following restrictions must hold…"; Fisher, Fleissig and Serletis (2001, p. 62) write: "Adding-up … restrictions require that …"; Cranfield, Eales, Hertel and Preckel (2003, p. 357) write: "Adding-up is imposed with …" Barnett and Serletis (2008, p. 213) write: "…the resulting theoretical restrictions are…".

[2] But Barten, somewhat mysteriously, also wrote (1969, p. 16): "However, it is quite arbitrary as to which equation should be dropped, and to avoid any asymmetry it seems more appropriate to estimate the system in its complete formulation."

[3] Data are available upon request.

Table 1. Test results

Hypothesis	**Degrees of freedom**	**Chi squared test**	**Critical value 1 percent**
Overall restrictions	12	67.59	26.22
Adding-up	6	45.76	16.81
Slutsky symmetry given homogeneity	6	33.41	16.81

Table 2. Maximum likelihood estimates of model (17)

Parameter	Coeff.	Std. err.	P-value	95% confidence interval
$\alpha_{(\text{BREAD \& CEREALS})}$	0.168	0.013	<0.001	0.144 – 0.193
$\alpha_{(\text{MEAT})}$	0.003	0.024	0.892	-0.044 – 0.050
$\alpha_{(\text{BEVERAGE})}$	0.306	0.021	<0.001	0.265 – 0.347
$\alpha_{(\text{OTHER FOOD})}$	0.598	0.031	<0.001	0.537 – 0.659
$\beta_{(\text{BREAD})}$	-0.005	0.006	0.434	-0.017 – 0.007
$\beta_{(\text{MEAT})}$	0.013	0.012	0.296	-0.011 – 0.038
$\beta_{(\text{BEVERAGE})}$	0.004	0.012	0.706	-0.018 – 0.027
$\beta_{(\text{OTHER FOOD})}$	-0.02	0.018	0.279	-0.055 – 0.016
$\gamma_{(\text{BR./BR.})}$	0.085	0.006	<0.001	0.072 – 0.097
$\gamma_{(\text{BR./ME.})}$	-0.033	0.006	<0.001	-0.045 – -0.021
$\gamma_{(\text{BR./BE.})}$	-0.011	0.006	0.057	-0.021 – 0.000
$\gamma_{(\text{BR./OT.})}$	-0.052	0.011	<0.001	-0.073 – -0.031
$\gamma_{(\text{ME./ME.})}$	0.198	0.016	<0.001	0.166 – 0.230
$\gamma_{(\text{ME./BE.})}$	-0.035	0.010	0.001	-0.055 – -0.015
$\gamma_{(\text{ME./OT.})}$	-0.103	0.019	<0.001	-0.140 – -0.066
$\gamma_{(\text{BE./BE.})}$	0.116	0.009	<0.001	0.097 – 0.134
$\gamma_{(\text{BE./OT.})}$	0.031	0.015	0.034	0.002 – 0.061
$\gamma_{(\text{OT./OT.})}$	0.233	0.025	<0.001	0.185 – 0.281

log-likelihood = 773.39; n = 119