Department of Economics and Finance

Chair of Applied Statistics and Econometrics

An AIDS model of consumer demand for meat in the United States

SUPERVISOR:

Prof. Saul Lach

CANDIDATE:

Francesco Casalena

Matr. **211841**

ACADEMIC YEAR 2018/2019

Index

1. INTRODUCTION

The Almost Ideal Demand System (AIDS) elaborated by Deaton and Muellbauer (1980), is by far the most widespread econometric model used for investigating household consumption behaviour. As a matter of fact, since its development it has given rise to a huge and ever-growing literature[1]. The model's outstanding popularity can be easily explained by its ability to compound all the standard properties from microeconomics with a desirable level of simplicity in estimation and restriction testing.

Nevertheless, despite all the aforementioned positive aspects, the AIDS is by no means flawless. In fact, in order to solve the intrinsic nonlinearity of the model, Deaton and Muellbauer proposed a linear approximation (henceforth LA-AIDS), which however leads to potentially disruptive bias. Thus, an important part of the literature has insofar focussed on finding what version of the AIDS model provides the best (bias-minimising) estimates of the real demand system. In other words, we could say that the model's name is very appropriate: such set of equations is indeed *almost ideal*, since, up to today, no model apparently performs better, although itself is far from being perfect.

In this paper I am estimating a small model of the retail demand for meat in the United States, using a sample that ranges from 1980 to 2017. After briefly exposing the models that preceded the AIDS, I am showing in detail the structure of Deaton and Muellbauer's model, including the analytical derivation of the core equations and all the pitfalls associated with its estimation. Eventually, I proceed with the implementation: two different versions of the AIDS are estimated on the available data. On each model, common misspecification tests are performed. Due to the still unsettled dispute on which version is the most efficient, the tracking performance of each of the two models is compared and contrasted.

Finally, the last section summarises the results obtained. Possible internal and external criticisms to my inference are identified and discussed, while a short ending note illustrates the methodology behind the modelling strategy.

[1] For a survey of the empirical analyses performed since 1980 to 1996 see Karagiannis et al. (2000).

2. MODELLING HOUSEHOLD CONSUMPTION

2.1. Historical background

Following the mathematical axiomatisation of microeconomic theory and the technical development and refinement of statistical tools in the 1930's and 1940's, applied research could not refrain from drawing her interest towards investigating individual household consumption behaviour. Under a historical point of view, empirical demand analysis of the post-war era looks like a natural development after the numerous and important results achieved by theoretical speculation and mathematical economics in the previous decades.

Indeed, with the publication of 'Value and Capital' by John Hicks in 1939, classical microeconomic theory underwent a process of thorough restructuring and unification, with the creation of a mathematically rigorous and logically consistent framework. The theoretical legacy of Alfred Marshall, Vilfredo Pareto and other contemporary economists was purged from all the cumbersome philosophical speculations (e. g. the assumption of cardinal utility) and packed up in order to create an easy-to-handle toolbox built in an axiomatic way. Furthermore, Hicks' volume provided an exhaustive treatment of comparative statics, which was for the first time applied systematically to general equilibrium. The mathematical analysis of equilibrium stability was improved, as such issue was tackled both under a static and dynamic point of view. Such results proved eventually invaluable, since they actually codified what is usually known today as standard textbook microeconomic theory.

The second stream of thought behind the development of empirical household consumption models is the debate that took place in the United Kingdom between Ronald Fisher and Karl Pearson during the 20's[2]. The introduction of the Maximum Likelihood estimator by Fisher in 1922 was a ground-breaking discovery, as it provided future researchers with a method which was technically rigorous and numerically easy to compute. In a series of papers spanning from 1912 to 1934, Fisher effectively laid the ground for modern statistical inference, as he explicitly defined a series of concepts, such as likelihood, efficiency, sufficiency and information, which still today play a pivotal role in econometrics.

Furthermore, after the overwhelming theoretical development that took place between the end of the 19th and the beginning of the 20th century, economic science started moving

[2] For a detailed account of the history see Aldrich (1997).

her interest towards empirical analysis, and the foundation of the Econometric Society in 1931 marked a clear step in this direction. Nevertheless, Fisher's contributions to statistical inference proved essential but by no means sufficient to allow economic theory to confront herself with reality. In fact, as Ragnar Frisch explicitly declared in the Editorial of the first issue of *Econometrica* in January 1933:

"[…] Econometrics is by no means as economic statistics. Nor is it identical with what we call general economic theory, although a considerable portion of this theory has definitely a quantitative character. Nor should be econometrics be taken as synonymous with the application of mathematics to economics. Experience has shown that each of these three viewpoints, that of statistics, economic theory, and mathematics, is a necessary, but not by itself a sufficient, condition for the real understanding of the quantitative relations in modern economic life. It is the *unification*[3] of all three that is powerful. And it is this unification that constitutes econometrics"

The union of high speculation with raw data entails a good deal of methodological issues which need to be tackled, as well as the full awareness on behalf of the researcher about the theoretical underpinnings and the ultimate scope of her investigation. Mathematical formalism and refined inferential techniques constitute essential tools, but alone they cannot solve the task of providing a reliable portrait of a complex and manifold reality under a numeric form. Whence the need to provide a stylised picture of the historical background in which the AIDS was developed, as well as of the models that were built before it.

2.2. Basic results from microeconomic theory

Before starting the discussion of the models that logically and chronologically preceded the AIDS, it is profitable to recall and summarise some basic properties derived from microeconomics that should be kept in mind while performing inference on consumption data. In fact, such results will be recurrent throughout the whole paper and a clear exposition of the basics is an essential starting point of the discussion that will take place in the following chapters.

If, as Lionel Robbins famously stated, "Economics is the science studying the relationship between ends and scarce means which have alternative uses", then, in order to

[3] Italics in the original.

analyse individual behaviour, it is necessary to start from an arbitrary optimisation problem. For example:

$$f(x_1, x_2, \lambda) = U(x_1, x_2) - \lambda(m - x_1p_1 - x_2p_2) \qquad (2.1)$$

Where, following a common notation, a two-commodity utility function U is maximised subject to some budget constraint m, and λ is the Lagrange multiplier. The first-order conditions are met as:

$$\frac{\partial f}{\partial x_1} = \frac{\partial U}{\partial x_1} - \lambda p_1 = 0$$

$$\frac{\partial \mathrm{f}}{\partial \mathrm{x}_2} = \frac{\partial \mathrm{U}}{\partial \mathrm{x}_2} - \lambda \mathrm{p}_2 = 0 \qquad (2.2)$$

$$\frac{\partial f}{\partial \lambda} = m - x_1p_1 - x_2p_2 = 0$$

and they form a system of three equations, with three unknowns x_1, x_2 and λ, and three exogenous parameters m, p_1 and p_2. The solution of such system clearly yields the equilibrium values of consumption. Furthermore, the Jacobian matrix of the system, i.e.:

$$\begin{bmatrix} \frac{\partial U^2}{\partial^2 x_1} & \frac{\partial U^2}{\partial x_1\,\partial x_2} & -p_1 \\ \frac{\partial U^2}{\partial x_2\,\partial x_1} & \frac{\partial U^2}{\partial^2 x_2} & -p_2 \\ -p_1 & -p_2 & 0 \end{bmatrix} \qquad (2.3)$$

is equivalent to the Hessian matrix of the function $f(x_1, x_2, \lambda)$. The study of the Hessian determinant, calculated by:

$$p_1p_2\left(\frac{\partial U^2}{\partial x_1\,\partial x_2} + \frac{\partial U^2}{\partial x_2\,\partial x_1} - \frac{\partial U^2}{\partial^2 x_2} - \frac{\partial U^2}{\partial^2 x_1}\right) \qquad (2.4)$$

yields information about the qualitative nature of the stationary point. In fact, since the prices are assumed to be strictly positive, the sign of the Hessian ultimately depends on the signs of the second and mixed derivatives of the utility function. Preferences are by assumption well-behaved, i.e. they generate indifference curves that are monotonic and convex w. r. t. the origin. This means that the first derivatives of the utility function are positive, while the second ones are negative, due to the principle of decreasing marginal utility. Since the first element (that is, in position a_{11}) of the Hessian matrix is negative, it is possible to conclude that the equilibrium point is indeed a maximum point.

If we regard the matrix (2.3) under the Jacobian point of view, by applying the implicit functions theorem[4], we know that, since the determinant is not zero, the functions:

$$x_1 = x_1(m, p_1, p_2)$$
$$x_2 = x_2(m, p_1, p_2) \qquad (2.5)$$
$$\lambda = \lambda(m, p_1, p_2)$$

exist, and the first two are the Marshallian (or Walrasian, or uncompensated) demand functions for the two commodities.

Marshallian demand functions retain two fundamental properties:

1) They are **homogeneous of degree zero**: the demanded bundle does not change for a scalar transformation of all prices and income, $x_i(p_i, p_j, m) = x_i(\alpha p_i, \alpha p_j, \alpha m)$, for any $\alpha > 0$.

2) They satisfy **Walras' law**, according to which if $n - 1$ markets are in equilibrium, then also the n^{th} one will be in equilibrium. This is equal to state that the consumer fully expends her wealth, hence total nominal income m and total money expenditure tend to coincide *ex ante*.

At this point, it is possible to consider a vector of Marshallian demand functions for n goods, where p is in turn a vector whose entries are the prices of the commodities. Instead, m denotes total disposable income.

$$x(p, m) = \begin{bmatrix} x_1(p, m) \\ \vdots \\ x_n(p, m) \end{bmatrix} \qquad (2.6)$$

The vector of the partial derivatives w. r. t. income of such demand functions is:

[4] See Samuelson (1947a, chap. 1-2) for a complete exposition in the realm of comparative statics.

$$D_m x(p,m) = \begin{bmatrix} \frac{\partial x_1(p,m)}{\partial m} \\ \vdots \\ \frac{\partial x_n(p,m)}{\partial m} \end{bmatrix} \tag{2.7}$$

While the following is the matrix whose entry $a_{i,j}$ represents the derivative of the demand function for good i with reference to the price of good j.

$$D_p x(p,m) \begin{bmatrix} \frac{\partial x_1(p,m)}{\partial p_1} & \cdots & \frac{\partial x_1(p,m)}{\partial p_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_n}{\partial p_1} & \cdots & \frac{\partial x_n(p,m)}{\partial p_n} \end{bmatrix} \tag{2.8}$$

Of course, normal goods have positive first partial derivatives with respect to income, while inferior goods will be demanded less as income increases. In a similar way, complementary goods are assumed to have a negative cross-price derivative (i.e. if goods i and j are generally consumed together, then an increase in the price of i will have *ceteris paribus* a negative impact on the demand for good j). Finally, a specular argument applies for the goods that are defined *substitutes*[5].

The Marshallian demand functions are derived from an optimisation problem: finding the maximum of the **indirect utility function**, that is, the function (2.1) describing the maximum level of utility (indifference curve) attainable, given some prices and income. If the prices vector is kept constant and the indirect utility function is inverted, it is possible to obtain the **expenditure** (or cost) **function**:

$$e(p,u) = m \tag{2.9}$$

It represents the minimum level of expenditure needed in order to achieve some level of utility. Logically, this implies that the expenditure function describes the minimisation problem of the budget constraint $p_1 x_1 + p_2 x_2 = m$ subject to some given indifference curve $U(x_1, x_2) = \overline{u}$.

Furthermore, if the expenditure function is differentiated w. r. t. prices, the **Hicksian** (or compensated) demand function is obtained. Its economic interpretation with respect to the

[5] For the sake of completeness, also Giffen goods (i.e. with a strictly positive own-price partial derivative) should be mentioned, however their importance in applied demand analysis seems to be negligible.

Marshallian one is pretty straightforward: it shows how quantities demanded change in response to changes in prices, while the *utility level* remains fixed. In other words, how the consumer will change her bundle in order to remain on the same indifference curve. On the other hand, Marshallian functions keep *expenditure* constant. Formally, the two are related as:

$$h(p,u) = x\big(p, e(p,u)\big) \qquad (2.10)$$

Finally, it can also be shown that:

$$\frac{\partial x_i(p,m)}{\partial p_j} = \frac{\partial h_i(p,u)}{\partial p_j} - \frac{\partial x_i(p,m)}{\partial m} x_j(p,m) \qquad (2.11)$$

In words, the price partial derivative of the Marshallian demand is equal to the price derivative of the Hicksian demand less the expenditure derivative of the Marshallian demand times the Marshallian demand of the commodity w. r. t. whose price the first derivative is taken. This is equivalent to disentangle the price effect (the movement along the same indifference curve) and the income effect (the change in utility level, keeping expenditure m constant) resulting from a variation in the price of a commodity.

The eq. (2.11) is called the **Slutsky equation**. Its second component, when expressed in matrix form, $D_p h(p,u)$, is also called the **Slutsky** (or **substitution**) **matrix** and it corresponds to the Hessian matrix of the expenditure function.

At this point, having deployed all the necessary conceptual instruments, it is possible move on to the discussion of the econometric models that historically have been developed with the scope of quantifying retail demand relations.

2.3. The Stone and Rotterdam models

The first individual expenditure model was the Linear Expenditure System (LES), elaborated by Richard Stone in 1954, which directly builds upon the premises exposed in paragraph [2.1]. In fact, the author introduces his paper by pointing out the five goals of his research:

1. To derive a practical system of equations possessing analytical properties soundly derived from microeconomic theory.
2. To consider the statistical problems involved in applying such system of equations.

3. To analyse the pattern of household consumption in the UK in the years 1920-1938.
4. To compare his findings with the actual state of demand in 1900.
5. To compare the post-war structure of demand with what might be expected from the inter-war relationships under free-market conditions.

In order to achieve his multiple purposes, Stone writes down a simple expenditure equation:

$$\hat{p}q = \hat{p}\overline{q} + b(\mu - p'\overline{q}) \tag{2.12}$$

Where $\overline{q}$ is the vector of quantities to which consumers are "committed", while b is a vector of weights summing up to unity and $\mu = p'q$ denotes total expenditure.

Under such hypothesis, consumption of the i-th good equals a certain quantity of committed consumption $\overline{q}_i$ valued at current prices, plus a certain proportion of supernumerary income, designed by total expenditure μ less committed one $p'\overline{q}$. This amounts to a "common sense" approximation to a demand system. Consumers hence spend a predetermined amount of their income in acquiring the consumption vector $\overline{q}$ at current (exogenous) prices, and then distribute their residual income over the set of available commodities in certain proportion given by the elements of vector b.

The system (2.12) is built upon the Klein-Rubin constant-utility index of the cost of living[6] and it can be rewritten as:

$$q = \hat{p}^{-1}\{b\mu + (bi' - I)\hat{c}p\} \tag{2.13}$$

Where i is the unit vector, I the identity matrix and $c = -\overline{q}$. If the whole (vector) equation is premultiplied by $\hat{p}$, it yields a system where expenditures on individual commodities are expressed as linear functions of only total expenditure and prices.

The system (2.13) may well be regarded as the effective ancestor of the AIDS. In fact, on the one hand it is a "parsimonious" model, since the coefficients of the price ratios depend only on $2m - 1$ coefficients, rather than $m^2 - 1$, when m is the number of total commodities in the model. Furthermore, its generality allows it to be consistent with the three canonical conditions usually imposed on demand systems in theoretical work:

[6] Elaborated by Klein and Rubin (1947). For a thorough discussion of its analytical properties see Samuelson (1947b).

(a) **Additivity**: the sum of expenditures as given by the system coincides *ex ante* with total expenditure. Analytically:

$$p'q \equiv \mu \qquad (2.14a)$$

(b) **Homogeneity**: for each commodity the sum of the total income and price elasticities is identically null:

$$\hat{q}^{-1}(a\mu + Ap) \equiv 0 \qquad (2.14b)$$

Where a is a vector whose entry $a\_i \equiv \frac{\partial \hat{p}q}{\partial \mu}$ is the derivative of quantity with respect to income, while A is the matrix of quantity derivatives w. r. t. prices, and 0 is of course the null vector.

(c) **Symmetry**[7]: the substitution matrix is symmetric, i.e. it coincides with its transpose. In formal terms it may be written as:

$$S \equiv S' \equiv \hat{q}^{-1}(ai' + A\hat{q}^{-1}) \qquad (2.14c)$$

The estimation can be more easily performed by reparametrizing (2.13) in a slightly different fashion, that is:

$$\hat{p}q = b\mu + Bp \qquad (2.15)$$

The matrix B can be easily determined as $B = (bi' - I)\hat{c}$, that is, by the multiplicative interaction of the vector of behavioural coefficients b with the vector of predetermined quantities in a way that is independent of prices[8]. In the case when c is assumed to be the null vector, B reduces herself to the null matrix and estimation can be smoothly carried out by OLS, regressing $p_i q_i$ on μ for each of the m equations. If, on the other hand, the vector of the given quantities is not null, we incur in a simultaneous equation bias, since each element of c enters every equation.

[7] Often referred as 'Slutsky symmetry'.

[8] The complete mathematical passages have been omitted due to the lack of time and space, since they are not fundamental for a synthetic exposition of the model. The complete analytical proof can be found in Stone (1954).

In order to overcome such issue, Stone opts for an iterated procedure: a provisional value of b, say b^*, can be estimated with the method above. Then, such obtained value can be plugged in the vector equation:

$$y_t = X_t c + u_t \tag{2.16}$$

Where $y_t \equiv (\hat{p}_t q_t - b^* \mu_t)$, $X_t \equiv (b^* i' - I)$ and u_t is the usual vector of residuals, on which several assumptions can be made. Then, an initial value $c^* = (X'X)^{-1} X'y$ is estimated with the usual OLS formula. With this estimate, it is possible to form the equations:

$$w_t = Z_t b + v_t \tag{2.17}$$

Where, in a similar fashion, $w_t \equiv \hat{p}_t (q_t + c^*)$, $Z_t \equiv \left(\mu_t + c^{*'} p_t\right)$ and v_t residuals. Now, a second-stage (least squares) estimate of parameter b, say b^{**}, is obtained. Thus, in estimating b, as opposed to c, the commodity groups are treated one at time. Starting from these new estimates, it is possible to re-estimate c and replicate the process a few times until 'stable' estimates of the two vectors are eventually reached.

Let us now resort to the analysis of the Rotterdam model, owing to the contributions of Anton P. Barten (1964) and Henri Theil (1965). Barten's declared attempt is to fill in the gap between consumer theory and empirical analysis. Noticeably enough, a flaw of the LES is that it is not explicitly derived from the maximisation of some utility function. The extreme "common-sense" (or "rule of thumb") approach gives it an excessive degree of arbitrariness, no matter how much rigorous the statistical estimation technique might be. As a consequence, the Rotterdam model is developed as a generalisation of the seminal contributions set out by Frisch (1959) and Houthakker (1960). Their specification is in fact restricted to direct or indirect additivity: too severe a constraint for the level of aggregation of the data usually available to the researcher. Hence, as the title of the 1964 paper explicitly points out, such assumption is straightforwardly removed.

Here is the functional form:

$$\Delta \log q_i(t) = \alpha_i + \sum_{j=1}^{n} \left(\varepsilon_{ij} - \varphi \eta_i \eta_j w_j - \eta_i w_j\right) \Delta \log p_j(t) + \eta_i \Delta \log \mu(t) + u_i(t) \tag{2.18}$$

The equation[9] is linear in its first logarithmic differences, with the intercept α_i that accounts for an autonomous trend due to changes in tastes. The dependent variable is the log-difference of quantity purchased of commodity i, while the regressors are the log differences of prices and income μ. Finally, $u_i(t)$ are the residuals, on which the usual assumptions of (null-mean) normality, homoscedasticity and serial uncorrelatedness are made[10]. The regression coefficients, instead, are slightly more complicated. Also in this case, there is no important gain in deriving extensively all the equations; therefore, I am limiting myself to exposing the economic intuition behind the econometric specification.

Now, let $\eta = \mu\overline{q}^{-1}q_{\mu}$ denote the column vector of income elasticities, which is obtained by multiplying the vector of total expenditure μ by the inverse of the matrix $\overline{q}$ (which is an n times n diagonal matrix whose diagonal is the quantities vector q) and by the vector of derivatives of quantities w. r. t. income q_{μ}. Let also $E = \overline{q}^{-1}Q_p\overline{p}$ be the matrix of price elasticities, which is defined is an analogous way to the former: the multiplication of the inverse of a matrix with all zero entries except the trace, which is in turn the quantities vector, $\overline{q}$, times the matrix of the derivatives of quantities w. r. t. prices and the diagonal matrix whose trace is equal to the prices vector, $\overline{p}$. Both the preceding formulas are simply the transposition in the realm of vector and matrix algebra of the common definition of elasticity, namely the ratio between the derivative and the fraction with reference to the variable of interest (in this two cases income and prices).

Barten shows that the matrix of the total price elasticities can be broken out in three distinct components: $E = E_1 + E_2 + E_3$. They are:

1. The first is related to the **direct substitution effect**, i.e. the direct interaction of the different commodities in the utility function, and it can be written down by:

$$E_1 = \overline{q}^{-1}(\lambda U^{-1})p \qquad (2.19a)$$

And it is obtained by multiplying the prices vector p times the inverse of the matrix $\overline{q}$ and the inverse of U, which is the hessian matrix of the utility function, times λ, being the

[9] In the exposition I am using the original notation by Barten (1964).

[10] Actually, Barten in his paper makes the specific assumption that post-war residual autocovariance is half of the pre-1939 due to the low quality of pre-war data.

lagrangean multiplier in the constrained maximisation problem and which has the economic interpretation of marginal utility of income.

2. The second is related to an **'overall' substitution effect**, namely:

$$E_2 = -\varphi\eta q'_\mu \overline{p} \qquad (2.19b)$$

Where, in the usual notation, $\overline{p}$ is a square diagonal matrix with trace equal to the price vector p; q'_μ is the transpose of the vector of derivatives of quantities w. r. t. income; η is the income elasticities vector. Finally, $\varphi \equiv \frac{\lambda}{\mu\lambda_\mu}$ is the reciprocal of the income elasticity of λ[11].

3. The third is related to the **income effect**:

$$E_3 = -\left(\frac{1}{\mu}\right)\eta q'\overline{p} \qquad (2.19c)$$

Also in this case, the reciprocal of income is multiplied times its elasticities vector, the transpose of the quantities vector and the diagonal matrix $\overline{p}$.

At the end of the day, after pre-multiplying everything by the vector of budget shares with entries $w_i \equiv \frac{p_i q_i}{\mu}$, algebraic manipulation allows to rewrite (2.19a-c):

$$(E)_{i,j} \equiv (E_1)_{i,j} + (E_2)_{i,j} + (E_3)_{i,j} \equiv \varepsilon_{i,j} - \varphi\eta_i\eta_j w_j - \eta_i w_j \qquad (2.20)$$

Where $\varepsilon_{i,j} \equiv \varphi\eta_i$. This is an indeed useful result, since it shows that the regression parameters are nothing more than the (semi)-elasticities of demand. The regressors, in turn, are the log-differences of, respectively, prices and income, which approximate their infinitesimal variations.

Barten brings in a little more algebra in the picture, thus allowing himself to reach a higher degree of both theoretical and analytical precision, as ties with microeconomics result strengthened with respect to the previous specification. Barten's self-set goal of "filling in the gap between theory and inference" is quite satisfactorily achieved: this is the starting point for the Almost Ideal Demand System.

[11] Frisch (1959) calls it 'money flexibility' or 'flexibility of marginal utility of income'.

3. THE ALMOST IDEAL DEMAND SYSTEM

3.1. Structure and properties of Deaton and Muellbauer’s model

3.1.1. Innovative features

As already stated in the introduction, the exposition of the models that preceded the Almost Ideal Demand System is not an end in itself. Rather, it is indeed necessary in order to better grasp all the details that ensured the success of Deaton and Muellbauer’s model, as well as to understand what design an empirical demand system should have so to provide a reliable representation of the real world. We have seen so far that Stone’s system was actually pioneering, as it rigorously described the desirable microeconomic properties a demand system ought have: additivity, homogeneity and Slutsky symmetry. Nevertheless, it fell short of a fundamental characteristic, namely to be derived from an underlying utility function. Barten (1964) brilliantly overcomes such shortcoming, by directly specifying a first-order approximation of a demand system starting from the Hessian matrix of some utility function.

Yet, the Rotterdam model leaves room for further research. Although it is firmly grounded on mathematical utilitarianism, it still remains at too high a level of abstraction. There is a general utility function (on which the usual assumptions of monotonicity and concavity apply) which is maximised through the Lagrange multiplier and from which a set of equations describing a (first-order approximation) demand system is derived. The approach is substantially correct, but it might be further improved: what Deaton and Muellbauer (1980) actually do is nothing more than deriving such set of equations by simulating the behaviour of a single, rational, utility-maximising individual. They start - rather than from some too generic preference ordering- from a specific class of preferences which allows perfect aggregation over consumers.

Deaton and Muellbauer claim that their model has considerable advantages over the aforementioned specifications, as it holds simultaneously all the desirable properties already individually possessed by the LES and the Rotterdam model. In particular, the AIDS is declared to:

1. Give an arbitrary first-order approximation to any demand system.
2. Satisfy the axioms of choice exactly.
3. Aggregate exactly over consumers without resorting to parallel Engel curves.
4. Have a functional form consistent with known household-budget data.
5. Avoid the need for non-linear estimation.

6. Allow testing homogeneity and symmetry restrictions through linear constraints.

As it will be shown later, although the AIDS undoubtedly brings consumer demand analysis to a higher level, perhaps Deaton and Muellbauer were a little too optimistic about their model -at least as long as property (5) is concerned. However -it should be made clear since the beginning- the major pitfalls of the model are concerned with the statistical issue of the estimation, which is never a plain task in economics in general. Also for this reason, the results of my analysis in the last chapter will be critically discussed bearing in mind the intrinsically complex issue of econometric inference. Right now, nonetheless, the 'deterministic' properties of the AIDS are straightforwardly showed.

Owing to the theorems of Muellbauer (1975), the so-called *PIGLOG*[12] preferences permit the representation of market demands as if they were the outcome of the decisions by a rational representative consumer. While Barten optimises a generic utility function subject to the expenditure/income constraint, the preferences belonging to the *PIGLOG* class are represented via the **cost** or **expenditure function**:

$$log\, c(u,p) = (1-u)\, log\, a(p) + u\, log\, b(p) \qquad (3.1)$$

Provided that u almost always[13] lies between 0 (subsistence) and 1 (bliss), while the positive linearly homogeneous functions $a(p)$ and $b(p)$ define, respectively, the cost of subsistence and bliss.

Specific functional forms for $a(p)$ and $b(p)$ are chosen so that they are sufficiently flexible for algebraic manipulation. Furthermore, they should also have enough parameters that at any single point the partial derivatives of $log\, c(u,p)$ exist and can be put equal to the ones of an arbitrary cost function. That is:

$$log\, a(p) = \alpha_0 + \sum_k \alpha_k\, log\, p_k + \frac{1}{2}\sum_k \sum_j \gamma^*_{k,j}\, log\, p_k\, log\, p_j \qquad (3.2)$$

$$log\, b(p) = log\, a(p) + \beta_0 \prod_k p_k^{\beta_k} \qquad (3.3)$$

[12] Acronym of Price-Independent Generalized Logarithmic. See the appendix in Deaton and Muellbauer (1980) for a complete exposition of its properties.

[13] The only exceptions occur when $b(p)$ is "more concave" than $a(p)$: in this (not very interesting) case $c(u,p)$ is also concave for a range of $u > 1$. For a thorough discussion see the appendix in Deaton and Muellbauer (1980).

From this it follows directly that the complete AIDS cost function is[14]:

$$log\, c(u,p) = \alpha_0 + \sum_k \alpha_k \, log\, p_k + \frac{1}{2}\sum_k \sum_j \gamma^*_{i,j}\, log_{p_k}\, log\, p_j + u\beta_0 \prod_k p_k^{\beta_k} \quad (3.4)$$

Where α_i, β_i and $\gamma^*_{i,j}$ are some parameters. The cost function (and hence its natural logarithm) is linearly homogeneous in p, as long as the conditions $\sum_i \alpha_i = 1$ and $\sum_j \gamma^*_{j,k} = \sum_k \gamma^*_{k,j} = \sum_j \beta_j = 0$ are satisfied. The two functions $a(p)$ and $b(p)$ are indeed sufficiently flexible and differentiable and have enough parameters so that $log\, c(u,p)$ always exists, by exploiting the fact that utility is ordinal and hence a point can always be chosen in a way such that its second derivative w. r. t. utility can be set equal to zero.

As far as aggregation theory is concerned, here it is not worth spending too much time discussing it, since the reader can directly refer to the appendix in Deaton and Muellbauer (1980) and the further literature mentioned therein. Nevertheless, it should be stressed, a major innovation of the AIDS with reference to the previous specifications is that is built upon the maximising behaviour of an *individual* consumer, which then is aggregated. Still, in the aggregation process, it retains its particular individual features, as for instance the age, composition and tastes of each household.

Aggregation simply consists in summing up the individual demand system of each household, by weighting it with some index k_h, representing the individual features of each consumer or household. If all households were identical, then $k_h \equiv 1$ for any h. Then, of course, the share of aggregate expenditure on good i in the aggregate budget of all households (denoted by $\overline{w}_i$) is given by:

$$\frac{\sum_h p_i q_{i,h}}{\sum M_h} \equiv \frac{\sum_h M_h w_{i,h}}{\sum M_h} \quad (3.5)$$

[14] These are estimated using time series data. By following the original notation by Deaton and Muellbauer (1980) the subscript t is dropped in order to ease the text.

3.1.2. Analytical properties

By applying Shephard's lemma[15], according to which the price derivatives of the cost functions are the quantities demanded, and by pre-multiplying both sides of (3.4) by $\frac{p_i}{c(u,p)}$, the budget shares can be obtained:

$$\frac{\partial\, log\, c(u,p)}{\partial\, log\, p_i} = \frac{p_i q_i}{c(u,p)} = w_i \qquad (3.5)$$

Therefore, the budget shares obtained by differentiating (2.13), are a function of prices and utility:

$$w_i = \alpha_i + \sum_j \gamma_{i,j}\, log\, p_j + \beta_i u \beta_0 \prod_k p_k^{\beta_k} \qquad (3.6)$$

Where $\gamma_{i,j} = \frac{1}{2}\left(\gamma^*_{i,j} + \gamma^*_{j,i}\right)$. As it can be noticed, the procedure starting from *PIGLOG* preferences eventually yields a system of equations in which budget shares are function of utility and prices: by assuming a constant level of utility, a certain minimum expenditure is needed in order to achieve such utility level. Then, if such expenditure level is substituted in the original equation, the (3.6), we have a system of equation according to which the expenditure shares depend directly on the relative prices and on income (i.e. total expenditure). This is, *sic et simpliciter*, the philosophy of the AIDS. In contrast to the Rotterdam model, we do not have any more generic log-differences of some abstract utility function, rather we have gained a more mathematically flexible, and hence simpler, system of demand equations.

As briefly mentioned above, for a rational consumer total expenditure M corresponds to $c(u,p)$. Hence, such equality may be inverted in order to express (unobservable) utility as a function of only prices and expenditure. Through this algebraic transformation, it is possible to obtain the indirect utility function. Once this is applied to eq. (3.6), we obtain the final form of the AIDS, in which expenditure shares are a function of relative prices and real income. Analytically:

$$w_i = \alpha_0 + \sum_j \gamma_{i,j}\, log\, p_j + \beta_i\, log\, M/P \qquad (3.7)$$

The trans-log price index P is instead defined as:

[15] See Shephard (1970).

$$log\, P = \alpha_0 + \sum_k \alpha_k \, log\, p_k + \frac{1}{2}\sum_j \sum_k \gamma_{j,k} \, log\, p_j \, log\, p_k \quad (3.8)$$

This is the ready-to-estimate model. At a first glance, the abundance of parameters can be immediately noticed. In particular, if the system is made up by n equations (i.e. there are n commodities), there will be in total $n(n+2) = n^2 + 2n$ coefficients to be estimated. Indeed, there are n of them in the intercepts vector α_0, other n in the vector β of derivatives w. r. t. income, finally there is the square matrix with entries $\gamma_{i,j}$. For this reason, in order to have better estimates and reduce the computation costs, a few restrictions based on *a priori* information should be imposed. Luckily enough, however, all constraints regarding the economic properties of the demand system happen to be linear -and this is a definite advantage of the AIDS specification.

Deaton and Muellbauer suggest that any kind of intuition based on economic theory (in particular, regarding possible substitutability/complementarity and the signs of the coefficients) should be used to constrain the entries $\gamma_{i,j}$. Furthermore, there are at least three other ‘standard’ constraints: **additivity**, **homogeneity** and **symmetry**. The former:

$$\sum_{i=1}^{n} \alpha_i = 1; \qquad \sum_{i=1}^{n} \beta_i = 0; \qquad \sum_{i=1}^{n} \gamma_{i,j} = 0 \qquad (3.9a)$$

Is consistent with Walras’ law, i.e. additive preferences. This is an automatically fulfilled constraint, therefore a pivotal assumption underlying the AIDS is that expenditure shares always add up to unity. In contrast with Barten’s model, this might look like a limiting assumption. However, if enough data are available so to encompass the whole basket a household (or individual) may purchase, the AIDS yields in principle very precise estimates. The second:

$$\sum_j \gamma_{i,j} = 0 \qquad (3.9b)$$

Represents the property of homogeneity. According to this, the demand function, both expressed in terms of absolute quantities and in budget shares, is homogeneous of degree zero in prices. This condition ensures that there is no money illusion: if prices change all at the same rate, so to leave the structure of relative prices unvaried, consumption remains unaffected, *ceteris paribus*.

$$\gamma_{i,j} \equiv \gamma_{j,i} \qquad (3.9c)$$

The third constraint, instead, depicts the property of Slutsky symmetry, which directly flows from Shephard's lemma. For it to hold, it is sufficient that the matrix with entries $\gamma_{i,j}$ is symmetric. In other words, the parameter estimating the effect on consumption of commodity i of a change in the price of commodity j should be equal to the effect on consumption of commodity j of the same change in the price of commodity i.

Finally, the two fundamental properties of monotonicity and concavity (which characterise the so-called "well-behaved preferences") cannot be directly imposed on the parameters. Still, there is the possibility of checking for them, by testing *ex post* the estimated coefficients of the model. The monotonicity condition, i.e. that the expenditure function is strictly decreasing in prices, can be ensured by checking that -according to Shephard's lemma- that quantities and expenditure shares are non-negative.

On the other hand, in order to test for concavity of the expenditure function (i.e. that the maximisation solution is unique), it is necessary that the expenditure function exists. Therefore, a necessary condition for concavity to hold is that symmetry holds too and that the trans-log price index is used. If the requirements are met, then it suffices to prove that the Slutsky matrix negative semidefinite: its eigenvalues are all non-positive. Both these operations will be of course carried out during the estimation of my model, if the underlying constraints will not be rejected.

The Slutsky matrix is yielded by double-differentiation of the expenditure function. Its entry $s_{i,j}$ is given by:

$$\frac{\partial c(u,p)^2}{\partial p_i \, \partial p_j} = \frac{M}{p_i p_j} c_{i,j} \tag{3.10}$$

Where:

$$c_{i,j} \equiv \gamma_{i,j} + \beta_i \beta_j \, log \frac{M}{P} + w_i w_j - \delta_{i,j} w_i \tag{3.11}$$

Elasticities can also be derived with low algebraic effort. Price elasticities of Marshallian (i.e. uncompensated for the income effect) demand functions are derived from the indirect utility function. They are:

$$\Theta_{i,j} = \frac{\partial x_i}{\partial p_j} \frac{p_j}{x_i} = -\delta_{i,j} + \frac{\gamma_{i,j}}{w_i} - \frac{\beta_i}{w_i} \left(\alpha_j + \sum_k \gamma_{k,j} \, log \, p_k \right) \tag{3.12}$$

Where $\delta_{i,j}$ is Kronecker's delta: a two-variable, discrete function taking up value 1 if $i = j$ and zero otherwise. Conversely, Hicksian price elasticities are obtained from compensated demand functions. Mathematically, this is equivalent to inserting the income elasticity and the Marshallian price elasticity in the Slutsky equation:

$$\Theta^{*}_{i,j} = \Theta_{i,j} + \left(\eta_i w_j\right) \tag{3.13}$$

The expenditure/income elasticity η_i is in turn defined as:

$$\eta_i = \frac{\partial x_i}{\partial m}\frac{m}{x_i} = 1 + \frac{\beta_i}{w_i} \tag{3.14}$$

Where the notation is self-explanatory.

3.1.3. The relation with the Rotterdam model

The connections between the AIDS and Barten's model are rather intuitive. However, they can also be explored a little further. First of all, it is worth stressing that the 1980 model takes advantage of the results obtained by Muellbauer (1975) as far as aggregation theory is concerned. This is an invaluable advantage of the AIDS, since it allows for a higher degree of formal flexibility. On the other hand, the functional form of both specifications is quite similar, although the individual parameters happen to have different interpretations.

In fact, if eq. (3.7) is first-differenced, we can obtain:

$$\Delta w_i = \beta_i \Delta\, log\frac{M}{P} + \sum_j \gamma_{i,j}\Delta\, log\, p_j \tag{3.15}$$

Which is of course devoid of the intercept due to differentiation. Now, if the translog price index is substituted inside:

$$\Delta w_i = \beta_i \left\{\Delta\, log\, x - \sum_k \alpha_k \Delta\, log\, p_k - \frac{1}{2}\sum_k \sum_j \gamma_{k,j}\Delta\left(log\, p_j\, log\, p_k\right)\right\} + \sum_j \gamma_{i,j}\Delta\, log\, p_j \tag{3.16}$$

It is possible to obtain a specification that is very similar to the Rotterdam one. If the trans-log index is substituted with the first difference approximation of the Stone index $\sum_j w_j \Delta\, log\, p_j$ (and this may well happen sometimes, as it will be shown further on), then the r. h. s. of the previous equation becomes identical to the Rotterdam model's one in eq. (2.18), here rewritten as[16]:

$$w_i \Delta \, log \, q_i = b_i \{\Delta \, log \, x - \sum w_k \Delta \, log \, p_k\} + \sum_j c_{i,j} \Delta \, log \, p_j \quad (3.15)$$

A striking difference between the two, on the other hand, is about the dependent variable, which in turn in Barten's model is the vector of budget shares multiplied by the changes in quantity demanded. Thus, if in the 1964 model $w_i \Delta \, log \, q_i$ is replaced by $w_i \Delta \, log \, w_i$, an addition of $w_i \Delta \, log \left(\frac{p_i}{M}\right)$, then the first-difference AIDS is obtained.

As already stated, both models carry the advantage of allowing to test homogeneity and symmetry restrictions with linear constraints. On the other hand, the negativity monotonicity condition directly applies to the matrix of the price effects in the Rotterdam, while this is not true for the AIDS. However, the 1980 model has the undoubtful advantage of being directly derived from an explicit set of demand functions (3.6) and preferences (3.1).

3.2. Econometric estimation

3.2.1. The linear approximation

Let us now come to what is often defined, strictly speaking, the econometric issue. No sooner the researcher embarks herself in the apparently harmless task of the estimation, she immediately stumbles into a problem: once (3.8) is substituted into (3.7), the model becomes highly nonlinear in the parameters. Deaton and Muellbauer explain that the equation yielded by substitution:

$$w_i = (\alpha_i - \beta_i \alpha_0) + \sum_j \gamma_{i,j} \, log \, p_j + \beta_i \left\{ log \, x - \sum_k \alpha_k \, log \, p_k - \frac{1}{2} \sum_k \sum_j \gamma_{k,jlog} \, p_k \, log \, p_j \right\} + u_i \quad (3.17)$$

is not very problematic in itself for Maximum Likelihood estimation, since the first-order conditions of the likelihood and score function are linear in α and γ given β and conversely, so that the model can be iteratively estimated, by concentrating on one parameter at time and keeping the others fixed. Nevertheless, the practical identification of

[16] The notation is by Deaton and Muellbauer, notice that the intercept has been omitted.

the intercept happens to pose some issues. Such term is in fact informed by the α_i's in eq. (3.8). In the 1980 paper it is argued that if the price indexes are too collinear, the trans-log price index won't be very sensitive to changes in its weights. As a result, changes in the intercept in (3.7) due to shifts in the α_0 can be offset in the α's with a negligible effect on $log\, P$. Luckily enough, however, economic theory can be of some help: the intercept term in eq. (3.7) has an immediate microeconomic interpretation, as it represents the expenditure required to achieve a minimum standard of living[17] when prices are unity. Therefore, choosing an *ad hoc* value will not be too difficult. As it will be shown later, this task has been pretty satisfactorily overcome in the subsequent literature.

Deaton and Muellbauer suggest, in any case, another estimation method: the so-called **Linear Approximation**. They correctly point out that, if the price index $log\, P$ were actually known, then the model would be completely linear in all parameters. In this case, there would be no problem in estimating the model equation-by-equation with OLS, which in turn would yield equivalent estimates of MLE if the residuals are normally distributed. Hence, at this point the overall problem reduces itself to the task of finding the price index that best approximates the trans-log one. Practically, since the 'true' value of the price index is not known, the researcher may exploit the information provided by the level of collinearity of the prices of the single commodities. Indeed, the authors propose to approximate P by some other index P^*, for instance the Stone index $P^* = \sum_j w_j P_j$. In the case when the approximate index is a linear transformation of the 'real' index, say $P \approx \varphi P^*$, then the estimable model will become:

$$w_i = (\alpha_i - \beta_i\, log\, \varphi) + \sum_j \gamma_{i,j}\, log\, p_j + \beta_i\, log\frac{M}{P^*} + u_i \qquad (3.18)$$

In fact, strictly monotonic transformations like the logarithmic one do not affect the properties of price indexes. In their first paper, Deaton and Muellbauer estimate the model with both methods (3.17) and (3.18). As long as the prices are sufficiently collinear, they claim that the Stone index approximation provides a precise linearization of the 'true' model. No formal proof of this statement is supplied, but it hinges on the intuitive argument that the less autocorrelated the sequence of weights φ_t, the less the autocorrelated the regression residuals and hence the most unbiased and consistent the estimators. In the "best-

[17] Or indifference curve, by recalling the meaning of the cost function.

of-all" case, when φ_t has no variance, i.e. it is a sequence of identical constants $\varphi_t \equiv \varphi^* \ \forall \ t \in T$, the validity of the inference is assured.

In the opposite case, this problem becomes very disruptive. The sequence φ_t directly enters the residuals series u_t, while, at the same time, it is also part of the index $log\ P^*$. If the regressors and the residuals are correlated, assumption number 1 of the OLS estimator is violated: the coefficients estimated as $\hat{\beta} \equiv (X'X)^{-1}(X'y)$ will be distorted and will not asymptotically converge to their real values.

In the literature, this is a largely controversial issue. Alston *et alii* (1994) demonstrate via Monte Carlo experimentations that the claim by Deaton and Muellbauer (1980) is justified: high collinearity in the price indexes helps reducing the bias and approximates well the 'real' values of the AIDS. They test four different specifications of the same demand system: the first is the standard AIDS elasticity formula, as set out by Deaton and Muellbauer:

$$\varepsilon_{i,j} = -\delta_{i,j} + \frac{\gamma_{i,j}}{w_i} - \frac{\beta_i}{w_i}\left(\alpha_j + \sum_k \gamma_{k,j}\, log\, p_k\right) \qquad (3.19a)$$

The second, instead, is the standard LA with the Stone index:

$$\varepsilon_{i,j} = -\delta_{i,j} + \frac{\gamma_{i,j}}{w_i} - \frac{\beta_i}{w_i}\left(w_j + \sum_k w_k\, log\, p_k\left(\varepsilon_{k,j} + \delta_{k,j}\right)\right) \qquad (3.19b)$$

The third is a variation of the LA-AIDS formula elaborated by Chalfant (1987):

$$\varepsilon_{i,j} = -\delta_{i,j} + \frac{\gamma_{i,j}}{w_i} - \frac{\beta_i}{w_i} w_j \qquad (3.19c)$$

The last one, finally, is a LA formula elaborated by Eales and Unnevehr (1988):

$$\varepsilon_{i,j} = -\delta_{i,j} + \frac{\gamma_{i,j}}{w_i} \qquad (3.19d)$$

Their results show that formula (3.19c) performs best. On average, (3.19b) and (3.19c) perform better than (3.19a) and (3.19d). Instead, as the degree of collinearity of prices increases, the accuracy of all formulas decreases, in particular of (3.19d) and (3.19a). While the claim by Deaton and Muellbauer is substantially true, it is also a fact that higher collinearity hinders parameter estimation. Due to all these reasons, in the second model I am estimating the elasticities using formula (3.19c), as used by Chalfant (1987).

Furthermore, it should be noted, if the trans-log is approximated by the Stone index, simultaneity bias arises too: the budget shares appear both on the right and on the left-hand side of eq. (3.18). Logically, does causality run from prices to budget shares or in the opposite direction? In the economic literature[18] an often-adopted second-best choice is the use of the lagged expenditure shares in the Stone index, $P_t^* = \sum_j w_{j,t-1} P_{j,t}$. Otherwise, another suggested approach in the literature[19] is the Three Stages Least Squares Estimator, where obvious candidates as instrumental variables are all the (logged) prices and total nominal expenditure. Finally, another common sense choice is to pick up a different price index which eliminates the simultaneity problem, for example the linear analogue of the Laspeyres index, $log\, P_t^{Ls} = \sum_j w_{j,0}\, log\, P_{j,t}$. Also in this case, for a deeper discussion of the relevant literature the reader is redirected to the paper by Henningsen (2017).

3.2.2. The Iterated Linear Least Squares estimator

As no direct solution to the problems presented by the LA-AIDS specification seems to exist, the nonlinear AIDS is apparently self-suggesting. Under the theoretical point of view, the classical model can be consistently and efficiently estimated by Full-Information Maximum Likelihood (FIML). Practically, however, the statistical package *systemfit* of the R software cannot ensure reliable estimate, as they mainly have problems of non-convergence. At the moment, its function for nonlinear estimation, *nlsystemfit*, is still under development. For this reason, the AIDS coefficients are obtained via the iteration of linear estimations.

The procedure is simple: in the first step, the shares equations are estimated through linear techniques (e.g. OLS), while the trans-log price index is held fixed. Then, the trans-log price index is updated with the newly-estimated coefficients. These two steps are repeated indefinitely until the coefficients converge. As a matter of fact, this is a very old technique, since it identical in the substance to the one used by Stone (1954) and exposed in paragraph [2.3]. Furthermore, this procedure has been systematically used in empirical investigations during the 1990's, although the proof of the consistency of its estimates has only been supplied by Blundell and Robin (1999), which eventually called it the Iterated Linear Least Squares Estimator (ILLE).

[18] See for example Eales and Unnevehr (1988) or Blanciforti et al. (1986).
[19] For a partial discussion of the relevant literature the reader is referred to Henningsen (2017).

Often, first-step estimates of the coefficient to plug into the trans-log index are obtained by the LA-AIDS or by some other log-linear index, like the Paasche or the Laspeyres. Instead, as the trans-log price index is not a fixed regressor, since it depends on stochastic estimated coefficients, the standard coefficient variance-covariance matrix obtained by the last step of the iterated regression process will not yield the right standard errors estimates. Blundell and Robin derive the true one:

$$\sigma(\hat{\theta}) = \hat{J}\left(\hat{\Sigma} \otimes \left(G(\hat{\theta})' G(\hat{\theta})\right)\right)(\hat{J}')^{-1} \qquad (3.20)$$

The vector of all the estimated coefficients -$\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$- is denoted by $\hat{\theta}$. $\hat{\Sigma}$ is the dispersion matrix of the regression residuals. $G(\hat{\theta})$ is the matrix of the constructed regressors for a conditional linear regression, given the estimated coefficients. Finally, $\hat{J}$ is defined as:

$$\hat{J} = \left(I \otimes G(\hat{\theta})'\right)\nabla(\hat{\theta}) \qquad (3.21)$$

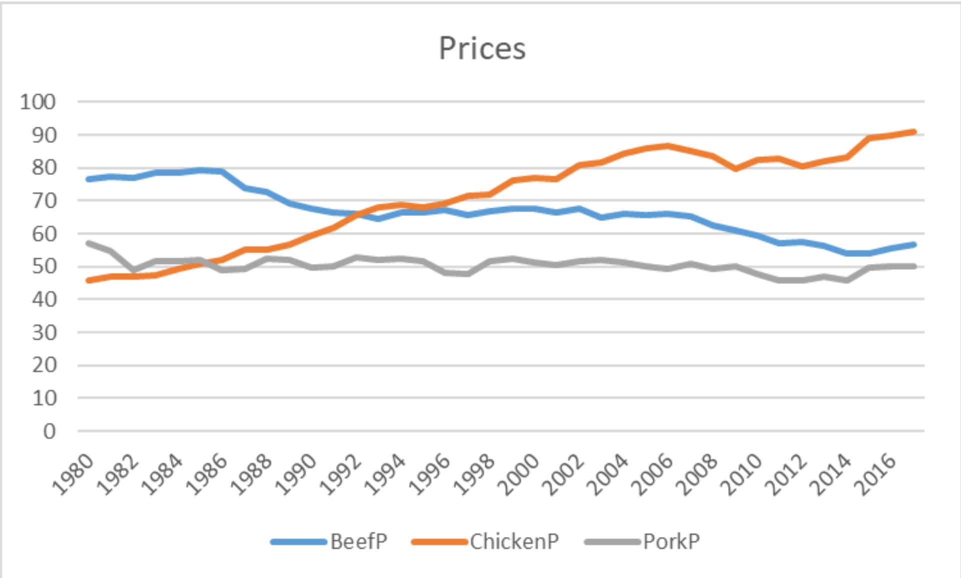
With the identity matrix I, the Kronecker product $\otimes$ and the Jacobian matrix of the expenditure shares w. r. t. the coefficients evaluated at their estimated values:

$$\nabla(\hat{\theta}) = \frac{\partial([I \otimes G(\hat{\theta})]\hat{\theta})}{\partial \hat{\theta}'} \qquad (3.22)$$

4. DATA AND RESULTS

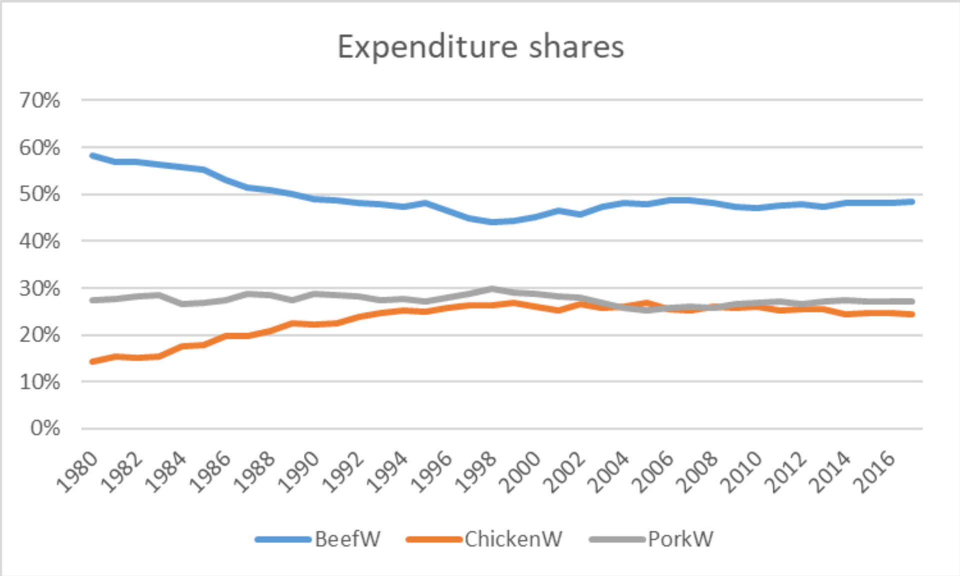
4.1. Data Sources

The data are freely available on the website of the U.S. Department of Agriculture[20]. They are yearly time series starting in 1980, with the last observation available of 2017. The three commodities represented in the demand system are Beef, Chicken and Pork. Quantity purchased is measured in pounds per capita at retail weight, i.e. excluding bones and non-edible components. Prices, on the other hand, are defined as the average retail price per pound of meat, expressed in United States Dollars. Total expenditure, as well as its relative shares, are simply obtained by summing up the money expenditure -price times quantity- on individual commodities.



Under the strictly statistical point of view, the average price (expressed in Dollar cents per pound) in the sample period is 66,6 c for beef, 70,8 c for chicken and 50,4 c for pork. Among the three series, the price of chicken clearly exhibits an upward trend. According to the data, in fact, the retail cost of a pound of chicken in 2017 is exactly the double of what it was in 1980. On the other hand, during the sample period the price of beef decreases of nearly 25%, where the most robust falls occur in the decades 1985-1995 and 2004-2014. Finally, the price of pork remains fluctuating stationarily about its sample average, except perhaps for a small temporary slide of circa 10% between 2008 and 2014.

[20] https://www.usda.gov/topics/data



The budget shares, after all, approximately tend to mirror the changes in the relative prices. The relative decrease of beef consumption with respect to chicken (probably triggered by an exogenous change in tastes?) leads to an increase of the price of the latter commodity relative to the former. Regarding pork, instead, its relative consumption remains stable over time -just like its price.

4.2. Inferential procedure

The steps of model estimation and selection in the following two paragraphs are pretty straightforward. In each one of those I am estimating a different specification of the AIDS. The first one is the classical nonlinear AIDS, whose coefficients are obtained by the ILLE and an 'optimal' intercept[21] is chosen. The starting point for the first-step estimates is the Laspeyres price index. The second model, instead, is the LA-AIDS. The trans-log index is approximated with the log-linear Laspeyres.

For every model, three different versions are estimated: one with the homogeneity and symmetry restrictions imposed, one only with the homogeneity restriction and, finally, a completely unrestricted version. It should be recalled here that the additivity property is satisfied *by definition* in the AIDS, hence it cannot be imposed (or removed) via algebraic restrictions on the parameters. The three versions of each model are then tested one against the other with the Likelihood Ratio test:

[21] The one that maximises the fit of the model, in terms of the likelihood value. For the details see Michalek and Keyzer (1992).

$$\lambda(x) = \frac{L(\theta^*;x)}{L(\hat{\theta};x)} \qquad (4.1)$$

The test statistic is made up by the ratio between the likelihood function maximised subject to a subset of the parameter space, θ^*, and the likelihood function evaluated at its maximum, $\hat{\theta}$. By definition, the statistic cannot exceed unity. The lower the support in favour of the null hypothesis, i.e. that the data-generating process is consistent with the vector of parameters θ^*, the lower the value of the ratio will be. The statistic is distributed with degrees of freedom equal to the difference in parameters of the two specifications.

At this stage, the 'best' version of the model in question is then studied more in detail. After the usual misspecification tests, the model's parameters are tested for significance and its elasticities are analysed and discussed. In addition, it is checked if the monotonicity (or even concavity, if symmetry is not rejected) condition is satisfied.

As the diagnostics is concerned, the scope of the testing procedure is to ensure that the residuals are well-behaved. This is indeed the most crucial part of the estimation process: for the model to be well specified, it is needed that its residuals do not contain any useful statistical information that can be used in the model to forecast the dependent variable. Formally, the regression residuals must satisfy three properties:

(i) **Serial independence**: the residual observed at time t (which may be regarded as the realisation of a random variable) should be statistically independent from the residual at time $t-1$. Formally, $E(u_t|u_{t-1}) = E(u_t) \; \forall \; t \in T$. If the residuals are normally distributed, this condition boils down to the absence of autocorrelation, since in the Gaussian distribution the moments greater than the second are all null.

(ii) **Homoscedasticity**: the variance of the residuals must be constant across the whole sample.

(iii) **Normality**: the residuals are normally distributed.

It should also be noted that the relevant condition for the estimates to be unbiased and consistent is (i). Hence, it should be absolutely met. As properties (ii) and (iii) are concerned, instead, failure to fulfil them does not invalidate the inference, since they only affect the standard errors of the estimated coefficients. According to the Gauss-Markov theorem, if the residuals follow a Gaussian distribution with constant variance, then the OLS is the most efficient estimator, i.e. the one with the least possible variance. However, should the last two requirements be unmet, this would not be a big problem, since s. e. can be computed in a consistent way. Therefore, the only true property that must be satisfied is the absence of serial dependence.

In general, a widespread indicator of the possible presence of autocorrelation in the residuals is the Durbin-Watson statistic, elaborated by Durbin and Watson (1958). It is simply calculated as:

$$\frac{\sum_{t=2}^{T}(u_t - u_{t-1})^2}{\sum_{t=1}^{T} u_t^2} \qquad (4.2)$$

This is equivalent to approximately calculating $2(1-\hat{\rho})$, where $\hat{\rho}$ is of course the sample correlation coefficient. Nevertheless, in the statistical package in R it is not possible to calculate the DW-stat for every individual equation. Therefore, I am investigating the autocorrelation by calculating the sample correlogram of the residual series. The graphs will be provided too in the following paragraphs.

As heteroscedasticity is concerned, instead, a common test to use is the one built by Breusch and Pagan (1979), although in the literature there are may different kinds of tests. Also in this case, in R it is not possible to perform post-estimation heteroscedasticity tests. On the other hand, this should not emasculate too much my inference since, as stated before, heteroscedasticity invalidates the standard errors of the coefficients, not the point estimates.

Finally, talking about normality, there is a great deal of different tests that can be implemented. In any case, they are all generally based on checking whether the sample moments of the residuals are consistent with those of a Gaussian distribution. Just to make an instance, the widespread Jarque-Bera parametric test has a statistic defined as:

$$\frac{n-k+1}{6}\left(S^2 + \frac{1}{4}(C-3)^2\right) \qquad (4.3)$$

Which is distributed as a Chi-Squared random variable with two degrees of freedom. The sample asymmetry coefficient is denoted by S, while C is the sample kurtosis.

The open-source environment that I am using for this computation, *Gretl*, automatically calculates all-together three different tests for normality: the Jarque-Bera, the Shapiro-Wilk and the Lilliefors test. Without discussing the structure of each, I am simply reporting the results in the appropriate section.

Given these premises, it is now time to have a look at the results.

4.3. Model 1: *Iterated Linear Least Squares*

This is the standard nonlinear model. Its estimates are yielded by the standard procedure set out by Blundell and Robin (1999). The optimal intercept is instead chosen using the procedure by Michalek and Keyzer (1992). Three different versions of this model are estimated. The first one bears the homogeneity and symmetry restrictions, the second one only homogeneity, while the third has no restrictions imposed on. They are tested one against the other to see what constraints are most supported by the data.

The results of the LR test are reported in the tab below: the models indicated in the row entries of the tab (the numerators) are tested against the models indicated in the columns of the tab (the denominators). The entries at the interceptions of rows and columns display the p-values, along with their significance levels. The latter are indicated with a common notation: three stars (***) for 0.01, two stars (**) for 0.05 and one star (*) for 0.1.

	Homogeneity	Unrestricted
Restricted	1.580e-15 (***)	1.942e-15 (***)
Homogeneity	-	0.0177 (**)

The results appear to be pretty clear-cut: the restricted specification is strongly rejected in favour of the one with the only homogeneity constraint imposed. In a similar fashion, the restricted model is clearly rejected also against the one with no restrictions imposed. The test also suggests that we should reject the model with the homogeneity restriction against the unrestricted one. However, in this case the statistic is significant only at 5% confidence level. We might assume that this result is probably imputable to the fact that the real data-generating process satisfies indeed the homogeneity restriction. Nevertheless, in my analysis many other commodities that might have some substitutability or complementarity relation with meat have been omitted. This might probably bias towards rejection the hypothesis of homogeneity.

In any case, *given the data available*, the LR tests indicate that the likelihood-maximising model, i.e. the one more consistent with the information provided by the data, is the unrestricted one. Although I am aware of the possible shortcomings of my choice, I will carry on analysing further the model with no restrictions, regarding it the most accurate description of reality. The following tabs summarise, respectively, the R^2 for each equation of estimated budget shares and quantities demanded and the coefficients of alpha's and beta's (standard errors in bracket), together with their p-values.

	R^2 shares	R^2 quantities
Beef:	0.9150982	0.9958739
Chicken:	0.9638942	0.9717371
Pork:	0.3726677	0.9884837

	alpha	p-value	**beta**	p-value
Beef:	-8.8192130	2.2e-16	0.1782839	2.2e-16
	(0.7646094)	***	(0.0123714)	***
Chicken:	6.8738102	2.2e-16	-0.1345425	2.2e-16
	(0.4505889)	***	(0.0072247)	***
Pork:	2.9454028	0.0002046	-0.0437413	0.0010967
	(0.7490651)	***	(0.0128101)	***

These are the estimated coefficients of the gamma's:

	Beef Price	Chicken Price	Pork Price
Beef:	-1.3675501	0.9845313	0.5606035
	(0.2152773)	(0.1249361)	(0.1682748)
Chicken:	1.0705975	-0.6769205	-0.4372269
	(0.1170867)	(0.1013974)	(0.0961277)
Pork:	0.2969526	-0.3076108	-0.1233766
	(0.1240065)	(0.0804027)	(0.0795636)

And these are their p-values:

	Beef Price	Chicken Price	Pork Price
Beef:	2.255e-08 ***	4.337e-11 ***	0.0014181 ***
Chicken:	2.419e-13 ***	6.082e-09 ***	2.374e-05 ***
Pork:	0.0194835 **	0.0002918 ***	0.1257639

The intercept of the model is **alpha_0** = -41.66667.

Monotonicity is fulfilled at 38 observations on 38 (100%).

Before passing on to the diagnostics, a couple of remarks are in order. While the first two equations seem to fit well the data observed, as their coefficients are all significant and carry the expected signs, the relatively worse performance in terms of fit of the Pork equation is pretty evident: the R^2 of the expenditure share equation is rather low (0.37, compared to 0.96 and 0.92). In addition, the γ_j coefficients have relatively lower p-values: the derivative w. r. t. beef price is significant only at 0.05 level, while the own-price derivative is not significant at all. This is a pretty surprising results, since it entails that the demand for pork does not depend on its price -and conversely.

It may be assumed that a possible explanation for this bizarre finding rests on the fact that (as shown in par. [4.1]) both the price and the expenditure share of pork have a rather low variance. By recalling that the OLS estimator with one regressor is yielded by $\widehat{\beta_\iota} = \frac{cov(y,x_i)}{var(x_i)}$, the intuition applies also in the multivariate case: insufficient variance in either the independent variable (or the dependent ones, or both) makes the parameter estimates very imprecise. Therefore, if they have an excessively large variance, the significance test is biased in favour of the non-rejection of the null hypothesis that $\hat{\beta} = 0$.

Bearing this in mind, it is possible to analyse the quality of the residuals of each equation. The following table shows the p-values of the test statistic for every equation, following each test:

	Jarque-Bera	Shapiro-Wilk	Lilliefors
Beef:	0.615137	0.749833	0.8
Chicken:	0.751262	0.875613	0.53
Pork:	0.432026	0.336367	0.25

In all cases we fail to reject the null hypothesis: the observed residuals are consistent with the realisation of a Gaussian random variable.

As autocorrelation is concerned, instead, things do not seem to go equally well. The table below shows the Partial Autocorrelation Function (PACF) of each equation, up to a maximum of 7 lags, and the respective significance level at each lag. These are the results:

LAG	Beef:	Chicken:	Pork:
1	0.5654 (***)	0.1808	0.6158 (***)
2	0.0642	0.0271	-0.2230
3	0.0428	0.2649	-0.0353
4	-0.3908 (**)	-0.1338	-0.0482
5	-0.0389	0.2438	-0.3620 (**)
6	-0.0653	-0.1251	-0.1889
7	0.1144	-0.0122	0.0331

The only equation whose residuals are totally uncorrelated is the one of the chicken shares. The other two, on the other hand, follow at least a first-order autoregressive scheme: the beef equation has a first-lag coefficient significant at 0.01 and a fourth-lag coefficient significant at 0.05. The pork equation has equally a first-lag coefficient significant at 0.01 and a fifth-lag coefficient significant at 0.05 level.

Therefore, the diagnostic tests inform us that all the results above should be taken with a grain of salt: it is likely that all the estimated coefficient are neither unbiased nor consistent. This shall not come as a surprise, given the numerous shortcomings of the analysis. In any case, I am concluding my inference by showing the estimated elasticities. Although they might be of difficult interpretation due to the potential bias of the parameters, it is worth reporting them down here for the sake of completeness.

Expenditure elasticities:

Beef:	Chicken:	Pork:
1.3669365	0.4381880	0.8407374

Hicksian (compensated) Price elasticities:

	Beef Price	Chicken Price	Pork Price
Beef:	-0.15269850	-0.13102721	0.6492233
Chicken:	0.09355214	0.08260366	-0.3580080
Pork:	0.18856045	0.15976912	-0.8363517

Taken literally, the expenditure elasticities imply that beef is a luxury good: for a certain percentage increase in income, the percentage increase in quantity demanded is more than proportional. On the other hand, chicken and pork look like being necessity goods, as their income elasticities are positive but below unity; for a given percentage increase income, the resulting percentage increase in demand will be relatively smaller. As the price ones are concerned, instead, the picture starts appearing more problematic: they are all very small in absolute value. Chicken's own-price elasticity is positive, although very close to zero. This would imply that chicken is a Giffen good, which seems unlikely. In addition, cross-price elasticities of beef to chicken price and of chicken to pork price are negative, thus implying some degree of complementarity between the two. Also this eventuality sounds rather counter-intuitive on logical grounds.

In any case, these are only the point estimates. It would be more interesting to know also the standard errors, so to test a few hypotheses. For example, it should be the case of testing if the beef income elasticity is greater than one and if chicken and pork income elasticities are less than one. At the same time, there is grounds for investigating whether the own-price elasticities are all significantly negative, while the cross-price ones lie strictly above zero. Unfortunately, the software package does not estimate the standard errors for the elasticities; in order to calculate them, I should compute the standard errors of the ratio of two random variables. A possibly viable way would be the linearization of the ratio and the application of the formula of the Delta method. However, that path would require too long calculations. Therefore, my analysis should stop here.

4.4. Model 2: *Linear Approximation - elasticities à la Chalfant (1987)*

Now, let us have a look at the second model: the one estimated using the linear approximation with the log-linear Laspeyres in place of the trans-log index. Although this model is possibly subject to all the biases exposed in the previous chapter, it should also be noted that the primary interest of my analysis is devoted mainly to the elasticities. These are calculated using the formula elaborated by Chalfant (1987), which should retrieve good estimates of the 'real' system.

The analytical procedure is identical to the one of the former paragraph: three versions of the LA-AIDS are estimated, each one differs from the other due to the constraints derived from microeconomics. The three versions are then tested to see which one the best is. Here are the results:

	Homogeneity	Unrestricted
Restricted	0.0011405 (***)	0.0001473 (***)
Homogeneity	-	0.0077785 (***)

In this case, the situation is clear: all the chi-square statistics are significant over the 1% confidence level. It is possible to reject the restricted and the model with the homogeneity constraint against the unrestricted one. The same argument of above applies: given the data available, the best model appears to be the one with no restrictions. These are its coefficients:

	R^2 shares	R^2 quantities
Beef:	0.9043119	0.9956788
Chicken:	0.9524978	0.9621421
Pork:	0.3892780	0.9890886

The alpha's and the beta's:

	alpha	p-value	**beta**	p-value
Beef:	-2.021765	1.737e-05	0.253904	1.931e-13
	(0.436241)	***	(0.027602)	***
Chicken:	1.694132	8.594e-07	-0.187741	5.320e-14
	(0.311815)	***	(0.019729)	***
Price:	1.327633	3.971e-05	-0.066163	0.0009233
	(0.301391)	***	(0.019069)	***

The gamma matrix:

	Beef Price	Chicken Price	Pork Price
Beef:	0.339020	-0.283258	0.157527
	(0.057110)	(0.022602)	(0.054028)
Chicken:	-0.211382	0.277851	-0.130980
	(0.040821)	(0.016155)	(0.038618)
Pork:	-0.127639	0.005407	-0.026547
	(0.039456)	(0.015615)	(0.037327)

And their p-values:

	Beef Price	Chicken Price	Pork Price
Beef:	1.194e-07 ***	2.2e-16 ***	0.0048439 ***
Chicken:	2.287e-06 ***	2.2e-16 ***	0.0011775 ***
Pork:	0.0019027 ***	0.7302461	0.4794634

In contrast to the AIDS estimated with ILLE, here the situation is even more clear-cut: in the $\gamma_{i,j}$ matrix pork has no significant derivatives w. r. t. the price of chicken and its own price. The lack of variance in the price of pork and its budget share has proved being a severe threat to the validity of my model.

Let us now come to the diagnostics. The p-values of the normality tests are reported in the next tab:

	Jarque-Bera	Shapiro-Wilk	Lilliefors
Beef:	0.928049	0.716962	0.77
Chicken:	0.679646	0.854026	0.41
Pork:	0.443262	0.373157	0.28

There are no problems in claiming that the residuals are normally distributed.

Here is instead the autocorrelation. Just like in the first instance, the last tab reports the value of the PACF for the seven first lags of the residuals, along with their significance level.

LAG	Beef:	Chicken:	Pork:
1	0.6171 (***)	0.4059 (**)	0.6063 (***)
2	0.1344	-0.0360	-0.0286
3	-0.2000	0.1311	-0.2175

4	0.0013	0.1083	0.0491
5	-0.0668	0.1887	-0.3611 (**)
6	-0.1662	-0.2141	-0.1590
7	-0.0373	-0.0932	0.1155

All the three equations have significant first-order autocorrelations. In addition, the residuals of the pork shares equation are also self-correlated at 5% significance level with their fifth lagged value. Hence, also in this case it is possible to conclude that the estimated coefficients are not very reliable. I am therefore ending with the exposition of the elasticities. These are the expenditure ones:

Beef:	Chicken:	Pork:
1.5139607	0.1957021	0.7572545

And these are the Hicksian:

	Beef Price	Chicken Price	Pork Price
Beef:	0.13557622	-0.2942138	0.5903834
Chicken:	-0.34162032	0.3521755	-0.2869225
Pork:	0.04683333	0.2316548	-0.8243387

The result is not radically different from the one obtained with the former model. The only macroscopic difference is that the own-price elasticity of beef is positive. Also in this case, this is a pretty weird result, since it would imply that beef is a Giffen good, which seem improbable. The identical argument of the previous paragraph applies: point estimates of elasticities are irrelevant as long as the variance is unknown. If the software does not compute the elasticities s. e., no hypothesis tests can take place.

5. CONCLUDING REMARKS

5.1. Overall model performance

Although in both cases the estimated coefficients and elasticities are approximately equal, the overall tracking performance of the models I have estimated is rather poor. The equations are misspecified: in both cases the residuals display significant first-order autocorrelation. This impairs the unbiasedness and consistency of the estimates. Furthermore, although the residuals are normally distributed, it is not possible to rule out the presence of heteroscedasticity. The software does not compute HC standard errors. In the case the residuals had not constant variance, the standard errors of the equation coefficients would be wrong. This means that, potentially, the significance tests performed on the estimates would be biased in favour of rejection.

In the table below, I am reporting the income and own-price elasticities for both of my models, while in the third column I am comparing them with the estimates by Moschini and Meilke (1989). They use the LA-AIDS over a quarterly dataset ranging from Jan 1967 to Oct 1987. The meaning of this comparison might be questionable, since the datasets of the two analyses overlap for only seven years. On the other hand, our two analyses are similar in the regressors used: Moschini and Meilke simply add fish to the commodities basket. In addition, the data source is identical: the statistics of the United States Department of Agriculture. Only the data regarding fish, being on an unpublished report, were unavailable for my model. Furthermore, in the 1989 paper the authors also test for a structural change in the model's parameters in 1976. I am only reporting the post-break elasticities, which are calculated on a dataset that has more in common with mine.

These are the expenditure ones:

	Model 1	Model 2	Moschini-Meilke
Beef:	1.367	1.514	1.394
Chicken:	0.438	0.196	0.211
Pork:	0.841	0.757	0.853

The coefficients of beef are all above 1 and also very close one to another. Also the elasticities of chicken and pork are rather similar. Overall, the income elasticities are consistent with economic theory and common sense. The first model, estimated via ILLE, returns estimates that are generally closer to the ones by Moschini and Meilke.

Generally speaking, for policy-making the only relevant elasticities are the Hicksian ones, since they are purged of the income effect. However, Moschini and Meilke only compute the Marshallian ones. Therefore, I am comparing my Marshallian estimates (which for the reason above were not reported in the previous paragraphs) with the ones of the 1989 paper. This comparison should be a kind of cross-check, in order to see if the values are approximately similar.

These are the own-price Marshallian elasticities:

	Model 1	Model 2	Moschini-Meilke
Beef:	-0.817	-0.612	-1.05
Chicken:	-0.022	0.3	-1.04
Pork:	-1.067	-1.03	-0.839

As predictable, the Marshallian elasticities are all smaller than the Hicksian ones due to the presence of income effect. Evidence shows that my estimates are still undervalued w. r. t. the ones by Moschini and Meilke. While pork has a very high absolute value, the other two are less than what expected. In particular chicken price elasticity in the first model is extremely close to zero, while in the second model is even positive. Therefore, these estimates are rather unreliable.

Finally, it must be stated that the third equation has a poor fit and some of its regressors are not significant. Arguably enough, the main culprit for this issue is the insufficient variance of the price and budget shares of pork meat consumption. Logically speaking, if from 1980 to last year prices and shares of pork consumption in the US have not varied significantly, we cannot know what it would happen if, starting from next year, the price of pork, say, increased by 30%. The lack of historical data prevents the inquiry from discovering the 'true' parameters of consumption in the US.

Apart from this flaw in the pork equation, it is reasonable to assume that the main bias these equations suffer from is the omitted variable one. There might be many commodities that have some degree of substitutability with those varieties of meat. Broadly speaking, when performing an analysis of individual consumer demand, it would be ideal to include as much goods as possible in the basket of the regressors: even if some commodities are not directly related to each other by consumption complementarity or substitutability, it cannot be assumed *a priori* that variations in their price will not influence via the income effect the demand for all the other commodities. In the real world, consumption bundles in indifference curves are made up by at least dozens of different goods: the more possible of them must be included in order to make the analysis valid.

5.2. Other possible criticisms

Up to now, I have discussed only the *internal* criticisms, i.e. the ones that pertain to my analysis. What about the *external* ones? Given that my analysis suffers from the shortcomings outlined above, what about the performance of the AIDS model *as such*? As a matter of fact, the secondary goal of my paper, other than attempting to estimate the demand for meat in the US, is to exemplify and discuss some of the methodological issues involved with econometric inference. I suppose that the application of the Almost Ideal Demand System provides a fruitful instance for all these problems. The goal is to find the best analytical description of the process that gave rise to the data available. In order to accomplish it, the functional form discussed above is deployed.

The first criticism that someone might levy against the AIDS is that the assumption of rational optimising behaviour is too restrictive: in the real life, people are potentially subject to mistakes while making their choices, hence the AIDS specification might not provide a reliable representation of reality. Yet, luckily, it is easy to rebut this argument. As Deaton and Muellbauer (1980) explain, their model does not stand or fall with the assumption of individual optimising behaviour: it is sufficient to assume that demand functions are differentiable w. r. t. prices and budget. If this condition holds, the AIDS is a first-order approximation to any demand system. This is a very important result, since it implies that, at least theoretically, the AIDS is an extremely powerful tool for applied analysis.

Therefore, the flaws of the model need to be tackled in the realm of statistics, not of microeconomic theory.

Probably, given a sufficient set of data, the ILLE-estimated AIDS and its Linear Approximation both provide good statistical representations of reality. Nevertheless, in the last years economic research has taken a few steps forward on this field. Many theoretical and empirical works have started analysing the AIDS further by exploiting the properties of time series analysis: just to make an instance, Karagiannis *et alii* (1997) have built an error-correction AIDS: they exploit the time series property of cointegration in order to disentangle short and long run interactions between prices and quantities demanded. Another interesting example is Moschini and Meilke (1989), who test the presence of structural breaks in the parameters of the demand for meat in the United States. Eclectic approaches just like those may prove very fruitful is tackling an issue that cannot be simply removed by brute calculations.

6. REFERENCES

Aldrich, J. (1997) "R. A. Fisher and the making of Maximum Likelihood 1912-1922", *Statistical Science*, **12**, 162-176.

Alston, J. M., Foster, A. K., Green, R. D. (1994) "Estimating elasticities with the Linear Approximate Almost Ideal Demand System: some Monte Carlo results", *The Review of Economics and Statistics*, **76**, 351-356.

Barten, A. P. (1964) "Consumer demand functions under conditions of almost additive preferences", *Econometrica*, **32**, 1-38.

Blanciforti, L., Green, R. (1983) "An Almost Ideal Demand System incorporating habits: an analysis of expenditure on food and aggregate commodity groups", *The Review of Economics and Statistics*, **65**, 511-515.

Blundell, R., Robin, J. M. (1999) "Estimation in large and disaggregated demand systems: an estimator for conditionally linear systems", *Journal of Applied Economics*, **14**, 209-232.

Breusch, T. S., Pagan, A. R. (1979) "A Simple Test for Heteroskedasticity and Random Coefficient Variation", *Econometrica*, **46**, 1287-1294.

Buse, A. (1998) "Testing homogeneity in the linearized Almost Ideal Demand System", *American Journal of Agricultural Economics*, **80**, 208-220.

Chalfant, J. A. (1987) "A globally flexible, Almost Ideal Demand System", *Journal of Business and Economic Statistics*, **5**, 233-242.

Deaton, A., Muellbauer, J. (1980) "An almost ideal demand system", *American Economic Review*, **70**, 312-326.

Durbin, J., Watson, G. (1958) "Testing for Serial Correlation in Least Squares Regression", *Biometrika*, **38**, 200-220.

Eales, J. S., Unnevehr, L. J. (1988) "Demand for beef and chicken products, separability and structural change", *American Journal of Agricultural Economics*, **70**, 521-532.

Fisher, R. A. (1922) "On the mathematical foundations of theoretical statistics", *Philosophical Transactions*, A, **222**, 309-368.

Frisch, R. (1933) "Editor's Note", *Econometrica*, **1**, 1-4.

Frisch, R. (1959) “A complete scheme for computing all direct and cross demand elasticities in a model with many sectors”, *Econometrica*, **27**, 177-196.

Henningsen, A. (2017), MicEconAids R package available at CRAN R project.

Hicks, J. R. (1939, 2nd ed. 1946) “Value and Capital: an inquiry into some fundamental principles of economic theory”, Oxford: Clarendon Press.

Houthakker, H. S. (1960) “Additive preferences”, *Econometrica*, **28**, 244-257.

Karagiannis, G., Katranidis, S., Velentzas, K. (2000) “An error correction almost ideal demand system for meat in Greece”, *Agricultural Economics*, **22**, 29-35.

Klein, L. R., Rubin, H. (1947) “A constant-utility index of the cost of living”, *The Review of Economic Studies*, **15**, 84-87.

Mas-Colell, A., Whinston, M. D., Green J. R. (1995) “Microeconomic Theory”, Oxford: University Press.

Michalek, J., Keyzer, M. A. (1992) “Estimation of a two-stage LES-AIDS consumer demand system for eight EC countries”, *European Review of Agricultural Economics*, **19**, 137-163.

Moschini, G., Meilke, K. D. (1989) “Modelling the pattern of structural change in U.S. meat demand”, *American Journal of Agricultural Economics*, **71**, 253-261.

Muellbauer, J. (1975) “Aggregation, income distribution and consumer demand”, *The Review of Economic Studies*, **62**, 525-543.

Robbins, L. (1932) “Essay on the nature and significance of Economic Science”, London: Macmillan.

Samuelson, P. A. (1947a) “Foundations of Economic Analysis”, Harvard University Press.

Samuelson, P. A. (1947b) “Some implications of ‘Linearity’”, *The Review of Economic Studies*, **15**, 88-90.

Screpanti, E., Zamagni, S. (1992, 2nd ed. 2005), “An outline of the history of economic thought”, Oxford: University Press.

Shephard, R. W., (1970) “Theory of cost and production function”, Princeton.

Spanos, A. (1986) “Statistical foundations of econometric modelling”, Cambridge: University Press.

Stock, J. H., Watson, M. (2015) “Introduction to econometrics: global edition”, London, Pearson Education.

Stone, R. (1954) “Linear expenditure systems and demand analysis: an application to the pattern of British demand”, *Economic Journal*, **64**, 511-527.

Theil, H. (1965) “The information approach to demand analysis”, *Econometrica*, **33**, 67-87.