Exact Affine Stone Index Demand System in **R**: The easi Package

Stéphane Hoareau
Université Laval

Guy Lacroix
Université Laval

Mirella Hoareau
Université Laval

Luca Tiberti
Université Laval

Abstract

easi is a package for R that enables the estimation of the Exact Affine Stone Index (EASI) demand system proposed by Pendakur and Lewbel. The EASI system is more flexible and easier to manipulate than traditional demand systems such as AIDS or QAIDS. It offers four major advantages: First, EASI budget shares are linear in parameters, conditional on real expenditures. Second, EASI demands are not constrained by the theoretic rank limit of Gorman. Third, unobserved preferences heterogeneity are taken into account via EASI error terms, which are equivalent to random utility parameters. Finally, EASI demands can be polynomials or splines of any order. Estimation of the EASI demand system has already been implement in Stata by Pendakur (2008). The **easi** package is more than a simple port of the Stata code. It offers a number of extensions within an unified framework : calculation of elasticities and equivalent income, simulations, selection of interaction terms, graphical representations of Engel curves with/without confidence intervals, *etc.*

Keywords:~households' expenditure survey analysis, EASI demand system, Engel curves, simulations, equivalent income.

1. Introduction

Most empirical analyses of consumer expenditure data rely on parametric demand models because they are relatively easy to estimate. Yet, it is well-known that these models are plagued with a number of empirical and theoretical shortcomings. Indeed, recent work has shown that many goods depict Engle curves that are highly nonlinear, even S-shaped (Blundell~Richard and Kristensen (2007)), a feature parametric models simply cannot account for. Furthermore, all parametric models face the so-called Gorman-type rank restrictons (Gorman (1981)). Finally, unobserved heterogeneity of preferences cannot readily be incorporated into most parametric models. Traditionally, the error terms are treated as random utility parameters (Bryan and Walker (1989), Daniel and Richter (1981), Donald and Matzkin (1998), Lewbel (2009), and Walter and Blundell (2004)).

The Exact Affine Stone Index (EASI) Demand System of Pendakur (2008) and Lewbel and Pendakur (2009) is a major breakthrough in the estimation of demand systems. It overcomes the aforementioned shortcomings while remaining "easi" to use. The main novelty of their approach is to express utility in terms of observed variables as derived from the Hicksian demand functions. Pendakur and Lewbel thus suggest we focus on what they refer to as *implicit* Marshallian demand functions. The latter are simply Hicksian demand functions in which the unobserved utility level is substituted out. Their *implicit* demand system avoids all the caveats mentionned above: It is linear in parameters, can easily incorporate unobserved heterogeneity, is not limited by the Gorman rank restrictions, it is capable of generating highly nonlinear Engle curves, and most of all, it is relatively "easi" to estimate.

Pendakur and Lewbel derive the theoretical properties of the EASI model and propose an estimation strategy (iterated 3SLS). In addition, they derive various budget/quantity elasticities. Finally, they illutrate the properties of the model using Canadian micro-data.

Currently, R offers two solutions for estimating demand systems. The **systemfit** package (Henningsen and Hamann (2011)) can estimate systems of linear and nonlinear equations using Ordinary Least Squares (OLS), Weighted Least Squares (WLS), Seemingly Unrelated Regressions (SUR), Two-Stage Least Squares (2SLS), Weighted Two-Stage Least Squares (W2SLS), and Three-Stage Least Squares (3SLS). The **micEconAids** package (Henningsen (2011)) focuses explicitely on the Almost Ideal Demand System (AIDS) suggested by Deaton and Muellbauer (1980). It is also based upon **systemfit**.

The **easi** package offers a unified framework within which the user can effortlessly estimate the EASI demand system as well as request additional statistical analyses. Thus in addition to the estimation of the model, the **easi** package can calculate predicted budget shares, generate graphical representations of Engel curves (with/without confidence intervals), calculate various budget shares (price, income, demographics) and quantity (price, income) elasticities, as well as calculate equivalent incomes. The package also simulates price changes, income changes, demographics changes and measures the impact on predicted budget shares and elasticities and on the shapes of Engel curves. The **easi** package is more flexible than the Stata code proposed by Pendakur (2008) in that the user can choose a subset of demographic variables to interact with prices and/or expenditure.

The paper is organized as follows: Section˜2 presents the EASI Demand System. The EASI model, the associated calculations and estimation method are presented in details. Section˜3 focuses on the calculations of the elasticities. We briefly recall the expressions for the budget shares presented in Lewbel and Pendakur (2009), after which we derive similar expressions for the quantity elasticities. Section ˜4 develops the expressions for the equivalent incomes in the context of EASI model. Section ˜5 provides a test for the local concavity of the EASI cost function. Section ˜6 presents the **easi** package structure. Section ˜7 illustrate the use of **easi** package with several examples. In particular, we replicate the estimation results of Lewbel and Pendakur (2009). Section˜8 concludes.

2. The EASI Demand System

Let $C(p, u, z, \epsilon)$ be a cost (expenditure) function, where p is the price vector, u is the utility level, z is the a vector of demographic variables which proxy observable preference heterogeneity. In additon, let ε be a vector of error terms which include unobservable preference heterogeneity. Hicksian compensated budget-shares functions can be derived using Shephard's lemma: $w = \omega(p,u,z,\varepsilon) = \nabla_p C(p,u,z,\epsilon)$. By expressing the indirect utility function in terms of g of w, p, x, z, the *implicit* utility function, y, is defined as $y = g(\omega(p,u,z,\varepsilon), p, x, z) = g(w,p,x,z)$. The *implicit* utility function depends only on observable data. Its closed-form expression is flexible, easily lends itself to empirical implementation, and does not depend on the utility have itself a closed-form expressions. The *implicit* Marshallian demand system is defined as $w = \omega(p,y,z,\varepsilon)$, which is simply the Hicksian demand system with y substituted in for u.

Lewbel and Pendakur (2009) refer to this class of cost functions as "Exact Affine Stone Index (EASI) cost functions", where y corresponds to an affine function of the Stone index deflated by the log nominal expenditures.

Lewbel and Pendakur propose the following cost function, which is particularly convenient for empirical implementation :

$$\begin{aligned} \ln C(p,y,z,\varepsilon) = y + \sum_{j=1}^{J} m^j(y,z)\ln p^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a^{jk}(z)\ln p^j \ln p^k \\ + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b^{jk}\ln p^j \ln p^k y + \sum_{j=1}^{J}\varepsilon_j \ln p^j \end{aligned} \tag{1}$$

Assuming there are J goods and T demographic variables, they propose the following parameterisation :

$$m^j(y,z) = \sum_{r=1}^{R} b_r^j y^r + \sum_{t=1}^{T} g_t^j z^t + \sum_{t=2}^{T} h_t^j z^t y \tag{2}$$

and $a^{jk}(z) = \sum_{t=1}^{T} a^{jkt} z_t$

The *implicit* Marshalian budget shares for each $j \in 1...J$ is then given by:

$$w^j = \sum_{r=1}^{R} b_r^j y^r + \sum_{t=1}^{T} g_t^j z_t + \sum_{k=1}^{T}\sum_{t=1}^{T} a_{jkt} z_t \ln p^k + \sum_{k=1}^{J} b_{jk}\ln p^k y + \sum_{t=2}^{T} h_t^j z_t y + \varepsilon_j \tag{3}$$

$$y = \frac{\ln x - \sum_{j=1}^{J} w_j \ln p^j + 1/2\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jkt} z_t \ln p^j \ln p^k}{1 - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b_{jk}\ln p^j \ln p^k} \tag{4}$$

These *implicit* Marshalian budget shares have several desirable properties. The linearity in parameters and additive error terms - namely the random utility parameters ε_j representing unobserved preferences - are certainly two of them. In addition, price effects can easily be interacted with expenditures and demographic characteristics. Engel curves can take virtually any shape through arbitrary high-order polynomials in log real expenditures. The demographic variables enter both through the intercept and the slopes of the log real-expenditures. Finally, EASI Engel curves for each good are almost completely unrestricted.

The strict monotonicity and concavity of the cost function are required in order to satisfy the others desirable properties of demand analysis - regularity, adding-up constraints, homogeneity and Slutsky symmetry.

Three methods of estimation are proposed in Pendakur (2008) and Lewbel and Pendakur (2009). All three take into account the heteroskedasticity of errors terms and the endogeneity of y. Parameters may be estimated by using Hansen's (1982) General Method of Moments or by homoskedastic nonlinear 3SLS. The estimators are consistent with heteroskedasticity but are only asymptotically efficient when the errors terms are homoskedastic. The authors recommend using an iterative linear 3SLS, which is a special case of a fixed-point based estimator considered by Dominitz and Sherman (2005).

3. Calculation of the elasticities

3.1. The Budget Shares Elasticities

Five types of budget shares elasticities are calculated in Pendakur (2008) and Lewbel and Pendakur (2009):

The semi elasticities of budget shares, Ψ, are given by :

$$\Psi = \frac{\partial \omega_j^i(p,y,z,\varepsilon)}{\partial \ln p^k} = \sum_{t=1}^{T} a^{jkt} z_t + b^{jk} y \tag{5}$$

The real expenditure semi-elasticities, $\aleph$, are given by :

$$\aleph = \frac{\partial \omega_j^i(p,y,z,\varepsilon)}{\partial y} = \sum_{r=1}^{R} b_r^j r y^{r-1} + \sum_{t=2}^{T} h_t^j z_t + \sum_{k=1}^{J} b^{jk} \ln p^k \tag{6}$$

The semi elasticities with respect to observable demographics, ζ, are given by :

$$\zeta = \frac{\partial \omega_j^i(p,y,z,\varepsilon)}{\partial z_t} = g_t^j + h_t^j y + \sum_{k=1}^{J} a^{jkt} \ln p^k \tag{7}$$

The compensated quantity derivatives with respect to prices, Γ ,are given by :

$$\Gamma = W^{-1}(\Psi + ww') \quad \text{where} \quad W = diag(w) \tag{8}$$

The compensated expenditures elasticities with respect to prices, S, are given by :

$$S = \Psi + ww' - W \tag{9}$$

3.2. The Quantities Elasticities

The elasticities for the quantities have been developed as part of the construction of the **easi** package.

Consider the EASI implicit marshallian demand system (3 and 4) and the following identity:

$$w^j = \frac{p_j Q_j}{x}, \tag{10}$$

where p_j is the nominal price of good j, Q_j is the amount of good j and x is the total expenditure.

Elasticity of good j with respect to the price of good i is given by:

$$\frac{\partial Q_j}{\partial p_i}\frac{p_i}{Q_j} = \eta_j^i \tag{11}$$

Therefore:

$$\frac{\partial Q_j}{\partial p_i} = \frac{\partial\left(\frac{xw_j}{p_j}\right)}{\partial p_i} + \frac{x}{p_j}\frac{\partial w_j}{\partial p_i} \tag{12}$$

Moreover :

$$\frac{\partial w_j}{\partial p_i} = A1 + A2 + A3 + A4, \tag{13}$$

with:

$$A1 = \frac{\partial(\sum_{r=1}^{R} b_r^j y^r)}{\partial p_i} \tag{14}$$

$$A2 = \frac{\partial(\sum_{t=2}^{T} h_t z_t y)}{\partial p_i} \tag{15}$$

$$A3 = \frac{\partial(\sum_{k=1}^{J}\sum_{t=1}^{T} a_{jkt} z_t \ln p_k)}{\partial p_i} \tag{16}$$

$$A4 = \frac{\partial(\sum_{k=1}^{J} b_{jk} \ln p_k y)}{\partial p_i} \tag{17}$$

Calculations of A_i give:

$$A1 = \frac{\partial y}{\partial P_i} \sum_{r=1}^{R} b_r^j y^{r-1} \tag{18}$$

$$A2 = \frac{\partial y}{\partial P_i} \sum_{t=2}^{T} h_t z_t \tag{19}$$

$$A3 = \sum_{t=1}^{T} a_{jkt} z_t \frac{1}{p_i} \tag{20}$$

$$A4 = \frac{1}{p_i} b_{ji} y + \frac{\partial y}{\partial p_i} \sum_{k=1}^{J} b_{jk} \ln p_k \tag{21}$$

Let $C = \sum_{r=1}^{R} b_r^j y^{r-1}$, $D = \sum_{t=2}^{T} h_t z_t$, $E = \sum_{t=1}^{T} a_{jkt} z_t$, $F = b_{ji} y$ and $G = \sum_{k=1}^{J} b_{jk} \ln p_k$

Moreover

$$y = \frac{u(p_i)}{v(p_i)}, \text{ where :} \tag{22}$$

$$u(p_i) = \ln x - \sum_{j=1}^{J} w_j \ln p^j + 1/2 \sum_{j=1}^{J} \sum_{k=1}^{J} \sum_{t=1}^{T} a_{jkt} z_t \ln p^j \ln p^k \text{ and} \tag{23}$$

$$v(p_i) = 1 - \frac{1}{2} \sum_{j=1}^{J} \sum_{k=1}^{J} b_{jk} \ln p^j \ln p^k \tag{24}$$

It follows that:

$$u'(p_i) = \frac{1}{P_i}(-w_i + \sum_{k=1}^{J} \sum_{t=1}^{T} a_{jkt} z_t \ln p^k) \text{ and} \tag{25}$$

$$v'(p_i) = -\frac{1}{P_i} b_{jk} \ln p^k \tag{26}$$

Furthermore :

$$\frac{\partial y}{\partial P_i} = \frac{u'(p_i)v(p_i) - u(p_i)v'(p_i)}{v^2(p_i)} \tag{27}$$

A little algebra allows us to write :

$$\frac{\partial y}{\partial p_i} = \frac{1}{p_i}B \tag{28}$$

Where

$$B = p_i \frac{u'(p_i)v(p_i) - u(p_i)v'(p_i)}{v^2(p_i)} \tag{29}$$

Hence

$$\frac{\partial Q_j}{\partial p_i} = \frac{\partial \left(\frac{xw_j}{p_j}\right)}{\partial p_i} + \frac{x}{p_j}\frac{1}{p_i}(B(C+D+G)+E+F) \tag{30}$$

Let $H = B(C+D+G)+E+F$.

Moreover

$$\frac{\partial \left(\frac{xw_j}{p_j}\right)}{\partial p_i} = \begin{cases} -\frac{xw_j}{p_j^2} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\frac{\partial Q_j}{\partial p_i} = \frac{Q_j}{p_i}\left(-1 + \frac{H}{w_j}\right) \tag{31}$$

Hence, letting η_j^i be the elasticity of good j with respect to the price of good i, we obtain :

$$\eta_j^i = -1(i == j) + \frac{H}{w_j} \tag{32}$$

Consider the EASI implicit marshallian demand system (3, 4) and the identity (10):

Let η_j^x be the elasticity of good j with respect to income. This elasticity is given by:

$$\frac{\partial Q_j}{\partial x}\frac{x}{Q_j} = \eta_j^x \tag{33}$$

So, we have :

$$\frac{\partial Q_j}{\partial x} = \frac{w_j}{p_j} + \frac{x}{p_j}\frac{\partial w_j}{\partial x} \tag{34}$$

Moreover :

$$\frac{\partial w_j}{\partial x} = \frac{1}{x}(C + D + G) \tag{35}$$

This allows us to write :

$$\frac{\partial Q_j}{\partial x} = \frac{1}{P_j}w_j + \frac{x}{p_j}\frac{1}{x}(C + D + G) \tag{36}$$

Hence, we obtain :

$$\eta_j^x = 1 + \frac{C + D + G}{w_j} \tag{37}$$

where $C = \sum_{r=1}^{R} b_r^j y^{r-1}$, $D = \sum_{t=2}^{T} h_t z_t$ and $G = \sum_{k=1}^{J} b_{jk} \ln p_k$.

4. Calculation of the equivalent income

The equivalent income is defined as the income level, $e_{c,h}$, that insures the utility levels are the same when evaluated at two prices vectors, *i.e.*:

$$v(p_c, x_{c,h}) = v(p_r, e_{c,h}), \tag{38}$$

where $v(.)$ is the indirect utility function, p_r is the reference price, and p_c is a different price vector. By inverting the indirect utility function, we obtain the equivalent income in terms of expenditure function: $e_{c,h} = e(p_r, p_c, x_{c,h})$, where $e_{c,h}$ is the equivalent income of household h living in stratum c, facing the price vector p_c, with a level of nominal income per capita (or per adult equivalent) $x_{c,h}$. The equivalent income $e_{c,h}$ is the level of income, at the reference price p_r, offers the same utility level than that obtained with the income level $x_{c,h}$ and the price system p_c. The function $e(p_r, p_c, x_{c,h})$ is increasing with respect to p_r and $x_{c,h}$, decreasing with p_c, concave and homogeneous of degree one with respect to the reference price, and has continuous first and second derivatives in all its arguments.

Consider the cost function (1) in the EASI class, where y is replaced by u . From (2) and (3), we have :

$$m^j(u, z) = w^j(p, u, z) - \sum_{k=1}^{J} a_{jk} \ln p^k - \sum_{k=1}^{J} b_{jk} \ln p^k u \tag{39}$$

By substituting (39) in (1), we have :

$$\begin{aligned}\ln C(p, u, z) = {} & u(1 - \sum_{j=1}^{J}\sum_{k=1}^{J} b_{jk} \ln p^j \ln p^k) + \sum_{j=1}^{J}\left(w^j(p, u, z) - \sum_{k=1}^{J} a_{jk} \ln p^k\right) \ln p^j \\ & + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jk} \ln p^j \ln p^k\end{aligned} \tag{40}$$

The contemporary situation is characterized by nominal total expenditures, $x_{c,h}$, and prices, p_c. This configuration achieves a level of utility $\overline{u}$:

$$\overline{u} = \frac{\ln x_{c,h} - \sum_{j=1}^{J} w^j \ln p_c^j + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jk} \ln p_c^j \ln p_c^k}{1 - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b_{jk} \ln p_c^j \ln p_c^k} \tag{41}$$

The reference or *ex ante* situation is characterized by nominal total expenditures equal to the equivalent income, $e_{c,h}$, and prices, p_r: This configuration also achieves a level of utility

$\overline{u}$. We can calculate this equivalent income $e_{c,h}$ by solving :

$$\ln C(p_r, u, z) = \ln e_{c,h} = \overline{u}(1 - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} b_{jk} \ln p_c^j \ln p_c^k) + \sum_{j=1}^{J}\left(w^j(p_r, u, z) - \sum_{k=1}^{J} a_{jk} \ln p_r^k\right) \ln p_r^j$$
$$+\frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jk} \ln p_r^j \ln p_r^k \tag{42}$$

By substituting (41) in (42), we obtain:

$$e_{c,h} = \exp\Bigg(\ln x_{c,h} + \textstyle\sum_{j=1}^{J} w^j \ln p_r^j - \sum_{j=1}^{J} w^j \ln p_c^j + + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jk} \ln p_c^j \ln p_c^k - \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{J} a_{jk} \ln p_r^j \ln p_r^k \Bigg) \tag{43}$$

5. Concavity of the cost function

The EASI demand system estimation and the calculation of the equivalent income assume that the cost function is concave. However, this concavity can be checked after estimation. Furthermore, it is well known that a semi-negative definite Hessian matrix is a necessary and sufficient condition to consider that the cost function is concave.

The Hessian matrix, H, is defined as:

$$H = \frac{\partial^2 C}{\partial p_j \partial p_k} \qquad \text{j,k } \in [1...J] \tag{44}$$

where C is the EASI cost function described in (1). By naming S the right hand side in the equation (1), we can write:

$$\ln C = S \Rightarrow C = \exp\{S\} \tag{45}$$

The Hessian matrix, H, is also equal to:

$$H = \frac{\partial^2 S}{\partial p_j \partial p_k} \exp S + \frac{\partial S}{\partial p_j}\frac{\partial S}{\partial p_k} \exp S \tag{46}$$

where:

$$\frac{\partial^2 S}{\partial p_j \partial p_k} = 2\frac{\partial^2 y}{\partial p_j \partial p_k} + \frac{\partial^2 S0}{\partial p_j \partial p_k} + \frac{\partial^2 S1}{\partial p_j \partial p_k} + \frac{\partial^2 S2}{\partial p_j \partial p_k} y + \frac{\partial S2}{\partial p_j}\frac{\partial y}{\partial p_k} + \frac{\partial S2}{\partial p_k}\frac{\partial y}{\partial p_j} \tag{47}$$

and

$$\frac{\partial S}{\partial p_j} = \frac{\partial y}{\partial p_j} + \frac{\partial S0}{\partial p_j} + \frac{\partial S1}{\partial p_j} + \frac{\partial S2}{\partial p_j} y + S2 \frac{\partial y}{\partial p_j} \tag{48}$$

A little algebra allows us to write:

$$S0 = \sum_{j=1}^{J} w^j \ln p_j \tag{49}$$

$$\frac{\partial S0}{\partial p_j} = \frac{w^j}{p_j} \tag{50}$$

$$\begin{aligned} \frac{\partial^2 S0}{\partial p_j \partial p_k} &= 0 \text{ if } j \neq k \\ &= -\frac{w^j}{p_j^2} \text{ otherwise} \end{aligned} \tag{51}$$

$$S1 = -1/2 \sum_{j=1}^{J} \sum_{k=1}^{J} \sum_{t=1}^{T} a_{jkt} z_t \ln p^j \ln p^k \tag{52}$$

$$\frac{\partial S1}{\partial p_j} = -\sum_{k=1}^{J} \sum_{t=1}^{T} \frac{a_{jkt}}{p_j} z_t \ln p^k \tag{53}$$

$$\frac{\partial^2 S1}{\partial p_j \partial p_k} = -\sum_{t=1}^{T} \frac{a_{jkt}}{p_j p_k} z_t \tag{54}$$

$$S2 = -1/2 \sum_{j=1}^{J} \sum_{k=1}^{J} b_{jk} \ln p^j \ln p^k \tag{55}$$

$$\frac{\partial S2}{\partial p_j} = -\sum_{k=1}^{J} \frac{b_{jk}}{p_j} \ln p^k \tag{56}$$

$$\frac{\partial^2 S2}{\partial p_j \partial p_k} = -\frac{b_{jk}}{p_j p_k} \tag{57}$$

6. Package structure

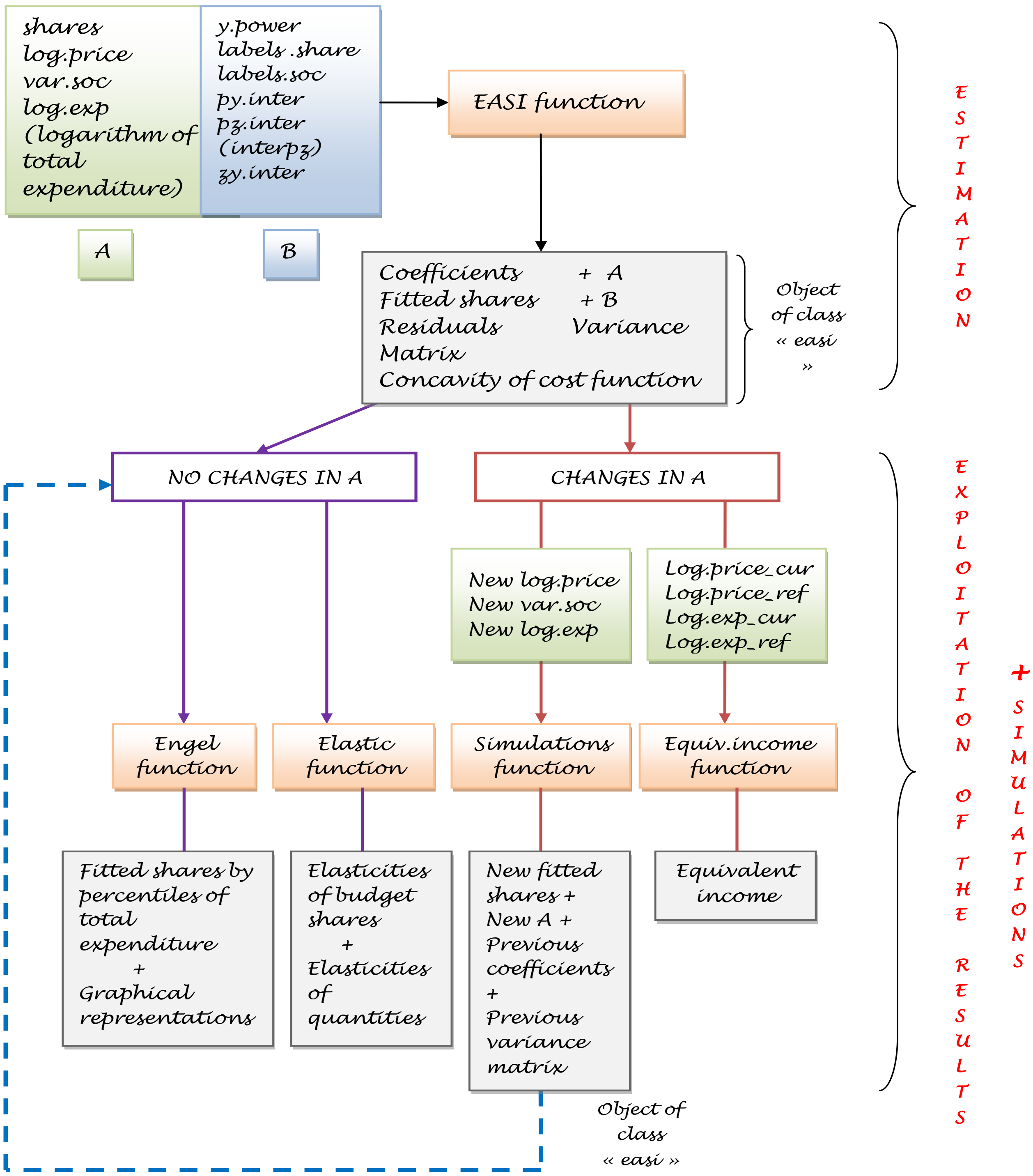


Figure 1: The Package Structure.

Figure 1 shows the **easi** package structure. The package has five main functions. The first function, namely the `easi` function, allows estimation of the model and generates nu,erous results as an object of class `easi`. Five methods were written to retrieve the results more easily: `coef` (parameter estimates), `predict` (matrix of the fitted budget shares), `residuals` (matrix of residuals), `summary` (summary of the estimation results), and `vcov` (covariance matrix of the paramter estimates). The `easi` function uses the **systemfit** package for the estimation procedure (iterated three stages least squares).

The results of the `easi` function (coefficients, variance matrix) can be used to compute Engel curves, elasticities, as well as to conduct simulations. More precisely, the `engel` function computes and draws the Engel Curves while the `elasticities` function calculates the elasticities of the budget shares and the elasticities of the quantities. The concavity of the cost function can be checked with the `concavity` function.

Likewise, two types of simulations are implemented within the **easi** package. Both require the user to specify new prices and / or new demographics and / or a new vector of total expenditure. The `simulations` generate new budget shares and `equiv.income` functions and equivalent income after above mentioned changes.

The result of `simulations` function is also an object of class `easi`. Therefore, the new elasticities and the new Engel Curves can be calculated after the simulations.

An internal function allows the calculation of intermediate blocks useful for generating the matrix of budget shares and the vector of utility implied. This function, namely `intermediate.blocs` is called by the `engel`, `elasticities`, `simulations` and `equiv.income` functions.

Finally, a database for examples and help is provided in the **easi** package. These data are those used by Lewbel and Pendakur (2009), namely the `hixdata`.

7. Examples

The **easi** package is loaded using:

```
R> library(easi)
```

To illustrate the use of the **easi** package, we use the `hixdata` data frame (See Lewbel and Pendakur (2009)).

```
R> data(hixdata)
```

Data consist of 4,847 observations of rental-tenure single-member canadian households that had positive expenditures on rent, recreation, and transportation (For details see Lewbel and Pendakur (2009)).

The covariates of this data are:

`obs`: number of observations

`sfoodh`: the budget share of food at home

`sfoodr`: the budget share of others foods

`srent`: the budget share of rent

`soper`: the budget share of household operations

`sfurn`: the budget share of household furnishing and equipment

`scloth`: the budget share of clothing

`stranop`: the budget share of transportation operations

`srecr`: the budget share of recreations

`spers`: the budget share of personal care

`pfoodh`: the logarithm of the price of food at home

`pfoodr`: the logarithm of the price of others foods

`prent`: the logarithm of the price of rent

`poper`: the logarithm of the price of household operations

`pfurn`: the logarithm of the price of household furnishing and equipment

`pcloth`: the logarithm of the price of clothing

`ptranop`: the logarithm of the price of transportation operations

`precr`: the logarithm of the price of recreations

`ppers`: the logarithm of the price of personal care

`log_y`: the logarithm of total expenditure

`age`: the person's age minus 40

`hsex`: the sex dummy equal to one for men

`carown`: a car-nonowner dummy equal to one if real gasoline expenditures (at 1986 gasoline prices) are less than 50 dollars

`time`: a time variable equal to the calendar year minus 1986

`tran`: a social assistance dummy equal to one of government transfers are greater than 10 percent of gross income

`wgt`: weighting variable

7.1. Estimation

The estimation is performed using the `easi` function, which has the following obligatory arguments: `shares` is the budget shares matrix (one observation by row and one item by column), `log.price` is the prices matrix (in logaritms) (one observation by row and one item by column), `var.soc` is the matrix of demographic variables and `log.exp` is the logaritm of total expenditure.

The user can customize the estimation with the following options. The first option - `y.power` - is an integer which corresponds to the highest desired power of `y` (implicit utility) in the system. `labels.share` and `labels.soc` are two strings which contain respectively the names of budget shares and the names of demographic variables. The following options - `py.inter`, `zy.inter` and `pz.inter` - are three logical variables which are each fixed to TRUE (FALSE otherwise) if the user wants respectively to enable the interaction between the price variables and `y`, the interaction between the demographic variables and `y`, the interaction between the prices and demographic variables. Finally, `interpz` is a vector which allows to choose which demographic variables have to be crossed with the price. For example, `interpz=c(3)` means that prices are crossed with the third demographic variable while `interpz = c (1:n)` means that prices are crossed with the first `n` demographic variables.

To illustrate the use of **easi** package, we reproduce the results of Lewbel and Pendakur (2009). We first estimate a simple EASI model, which we arbitrarily call `est`.

```
R> #****** Budget Shares Matrix *************
R> #     shares_HIX=hixdata[,2:10]
R> #****** Price Matrix (in logarithms) *****
R> #     log.price_HIX=hixdata[,11:19]
R> #****** Demographic matrix **********
R> #     var.soc_HIX=hixdata[,21:25]
R> #****** Logarithm of total expenditure ***
R> #****** (here divised by a price index) **
R> #     log.exp_HIX=hixdata[,20]
R> #****** Labels of demographic variables **
R> #     labels.soc <- c("age","hsex","carown","time","tran")
R> #****** Labels of budget shares **********
R> #     labels.share=c("food in","food out","rent","operations",
R> #       "furnishing","clothes","transport","recreation")


R> #est <- easi(shares=shares_HIX,log.price=log.price_HIX,var.soc=var.soc_HIX,
R>  #           y.power=5,log.exp=log.exp_HIX,labels.share=labels.share,
R>   #          labels.soc=labels.soc,py.inter=TRUE, zy.inter=TRUE,
R>    #         pz.inter=TRUE, interpz=c(1:ncol(var.soc_HIX)))
```

Several methods for an easier manipulation of the results. The `coef` method allows to recover the coefficients. Here, for more readability, we present only the first three lines of the coefficient matrix.

```
R> #head(coef(est),3)
```

The `vcov` method allows to recover the variance matrix. The covariance matrix is too large for printing here. Its size is given by:

```
R> #dim(vcov(est))
```

The `predict` method allows to recover the fitted budget shares:

```
R> #head(predict(est),3)
```

The `residuals` method allows to recover the residuals.

```
R> #head(residuals(est),3)
```

The `concavity` function allows to check the local concavity of the cost function.

```
R> #head(concavity(est))
```

Here, the cost function is concave on more than 90% of the observations.

7.2. Elasticities

The elasticities can be calculated with the `elasticities` function. The arguments of the function are the result of the `easi` function (an object of class `easi`, here `est`), the type of desired elasticities (between "price", "demographic" and "income") and a logical variable `sd` that indicates if standard deviations must be calculated.

The calculation of price elasticities are performed by:

```
R> #elastprice <- elastic(est,type="price",sd=TRUE)
```

The price elasticities of budget shares are recovered by:

```
R> #elastprice$EP[paste("p",labels.share,sep=""),
R>  #             paste("p",labels.share,sep="")]
```

The corresponding standard deviations are recovered by:

```
R> #elastprice$EP_SE[paste("p",labels.share,sep=""),
R>   #              paste("p",labels.share,sep="")]
```

The elasticities of quantities with respect to prices are recovered by:

```
R> #elastprice$ELASTPRICE[paste("p",labels.share,sep=""),
R> #                      paste("p",labels.share,sep="")]
```

The corresponding standard deviations are recovered by:

```
R> #elastprice$ELASTPRICE_SE[paste("p",labels.share,sep=""),
R> #                         paste("p",labels.share,sep="")]
```

The calculation of income elasticities are performed by:

```
R> #elastincome <- elastic(est,type="income",sd=FALSE)
```

The income elasticities of budget shares are recovered by:

```
R> #elastincome$ER[1,labels.share]
```

The elasticities of quantities with respect to income are recovered by:

```
R> #elastincome$ELASTINCOME[1,labels.share]
```

The calculation of demographic elasticities are performed by:

```
R> #elastdemographic <- elastic(est,type="demographics",sd=FALSE)
```

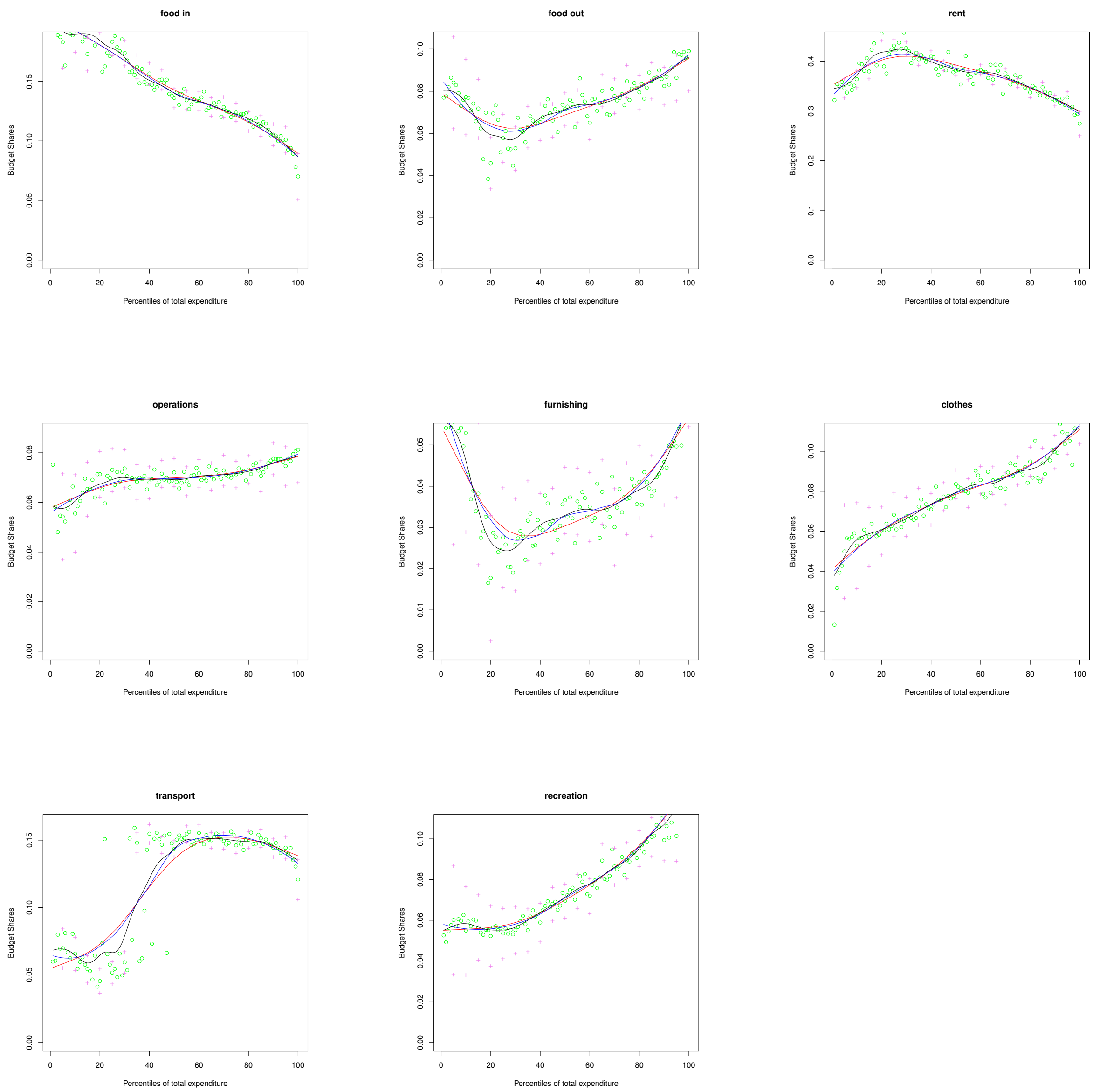
7.3. Engel curves

If one wants to calculate Engel curves, one can use the `engel` function, whose arguments are the result of the `easi` function (an object of class `easi`, here `est`), the name where the Engel curves must be graphically represented, and a logical variable `sd` that indicates if confidence intervals must be calculated and represented.

```
R> #eng1 <- engel(est,file="graph_engels_curves",sd=TRUE)
```

The figure 2 shows the Engel curves of the `est` estimation. Each green circle is the median of the budget share for the considered percentile of total expenditure described in abscissa. Magenta crosses delimit a confidence interval of 95%. Curves (black, blue and red) correspond to three increasing levels of smoothing.

Figure 2: The Engel Curves of easi estimation.



7.4. Simulations

easi provides routines to perform simulations. To illustrate, we decide to reproduce the simulation of Lewbel and Pendakur (2009), namely the estimated Engel curves from the model for a 40-year-old car-owning female in 1986 who didn't receive much government transfer income and having $\varepsilon = 0$.

This configuration implies that all prices (in logarithms) and all demographic variables are null.

```
R>  #    log.price_HIX.SIM1 <- log.price_HIX
R>  #   for (i in 1:ncol(log.price_HIX))
R>  #  log.price_HIX.SIM1[,i] <- 0
R>
R>  #    var.soc_HIX.SIM1 <- var.soc_HIX
R>  #    for (i in 1:ncol(var.soc_HIX))
R>  #    var.soc_HIX.SIM1[,i] <- 0
```

The `simulations` function allows to calculate the fitted values of budget shares after the previous changes in prices and demographics.

```
R>   #   sim <- simulations(est,log.price_new=log.price_HIX.SIM1,
R>   #          var.soc_new=var.soc_HIX.SIM1,log.exp_new=log.exp_HIX)
```

The corresponding Engel curves are calculated as previously.

```
R> #eng2 <- engel(sim,file="simeng",sd=TRUE)
```

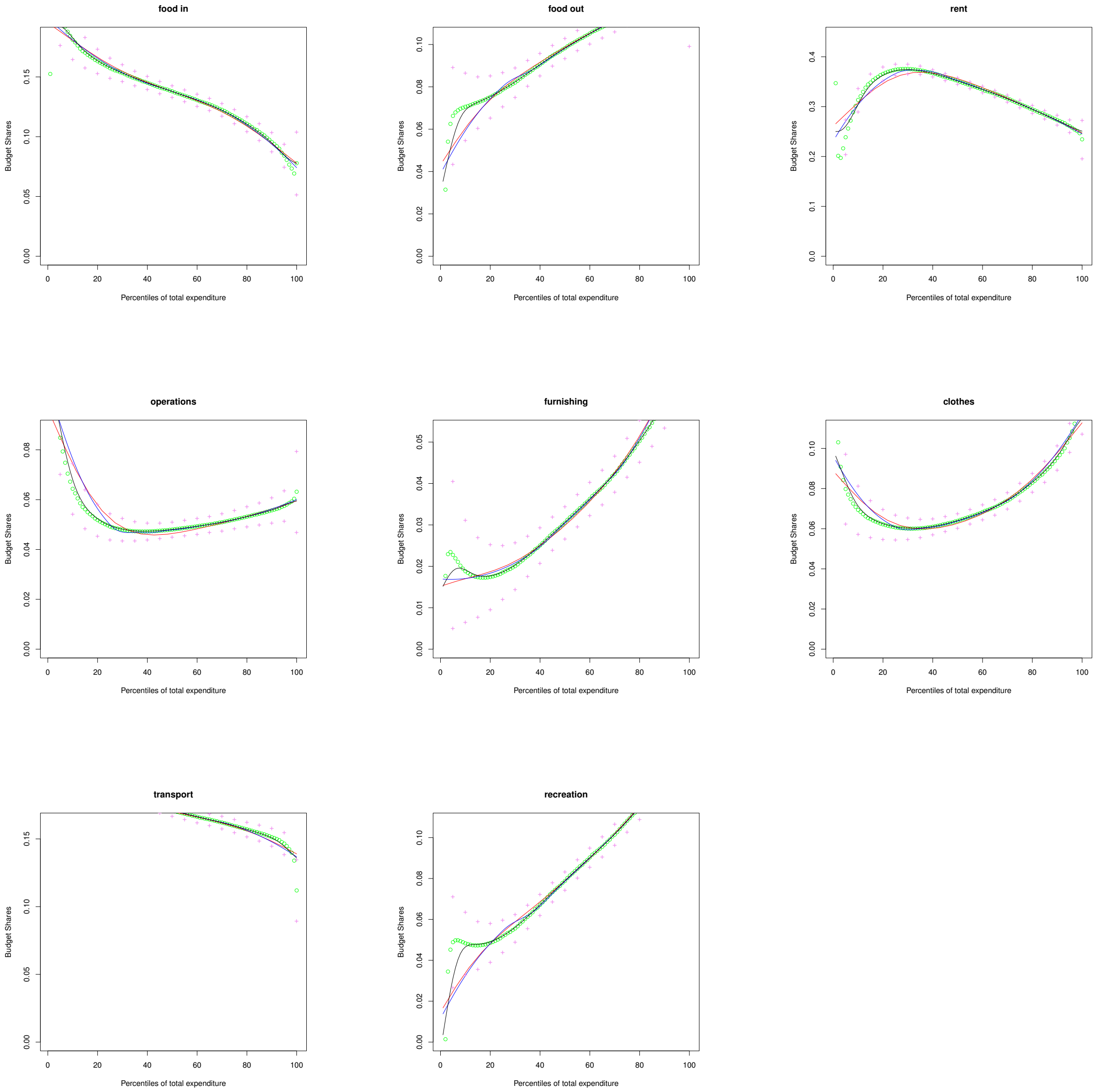
The figure 3 shows the Engel curves of the `sim` simulation.

7.5. Equivalent income

Our version of `hixdata` does not contain the variable ”total expenditure”. We propose to simulate instead an ”hybrid model” in order to illustrate the use of the `equiv.income` function:

This model is composed of five budget shares whose means are respectively equal to 0.25, 0.15, 0.20, 0.30 and 0.10.

Figure 3: The Engel Curves of easi simulation.




```
R> w1 <- rnorm(3000,mean=0.25,sd=0.05)
R> w2 <- rnorm(3000,mean=0.15,sd=0.05)
R> w3 <- rnorm(3000,mean=0.20,sd=0.05)
R> w4 <- rnorm(3000,mean=0.30,sd=0.05)
R> w5 <- 1-w1-w2-w3-w4
R> shares_SIM <- data.frame(w1,w2,w3,w4,w5)
```

We simulate five price vectors, whose means are respectively equal to 25, 15, 20, 30 and 10:

```
R> p1 <- log(rnorm(3000,mean=25,sd=3))
R> p2 <- log(rnorm(3000,mean=15,sd=2))
R> p3 <- log(rnorm(3000,mean=20,sd=3))
R> p4 <- log(rnorm(3000,mean=30,sd=4))
R> p5 <- log(rnorm(3000,mean=10,sd=1))
R> log.price_SIM <- data.frame(p1,p2,p3,p4,p5)
```

We simulate four demographics variable : V1, V3, V4 that are dummy variables, and V2 that take his values in $\mathbb{N}^+$.

```
R> V1 <- abs(round(rnorm(3000,mean=0.7,sd=0.2)))
R> V2 <- abs(round(rnorm(3000,mean=2,sd=1)))+1
R> V3 <- abs(round(rnorm(3000,mean=0.7,sd=0.2)))
R> V4 <- abs(round(rnorm(3000,mean=0.7,sd=0.2)))
R> var.soc_SIM <- data.frame(V1,V2,V3,V4)
```

Finally, we simulate a vector of total expenditure whose the average is 1200.

```
R>      log.exp_SIM <- log(rnorm(3000,mean=1200,sd=200))
```

The first step is to estimate the EASI model:

```
R>      est2 <- easi(shares=shares_SIM,log.price=log.price_SIM,
+                  var.soc=var.soc_SIM,log.exp=log.exp_SIM)
```

```
 *** Please wait during the creation of final instruments... ***
iteration =  1
crit_test =  0.0050925
iteration =  2
crit_test =  1.85e-05
iteration =  3
crit_test =  5e-07

 *** Creation of final instruments successfully completed... ***

 *** Please wait during the estimation... ***
iteration =  1
crit_test =  1.85e-05
iteration =  2
crit_test =  1.85e-05
```

```
iteration =  3
crit_test =  5e-07

 *** Estimation successfully completed ***
```

Let's consider the calculation of equivalent income after following changes : PRICE_SIM are multiplied by 1.4 between reference and current period while exp_SIM is only multiplied by 1.05 between reference and current period. This scenario thus corresponds a priori to a loss of purchasing power.

```
R>      log.price_SIM2 <- log(exp(log.price_SIM)*1.4)
R>      log.exp_SIM2 <- log(exp(log.exp_SIM)*1.05)
R>      equiv <- equiv.income(est2,log.exp_ref=log.exp_SIM,log.exp_cur=log.exp_SIM2,
+                  log.price_ref=log.price_SIM,log.price_cur=log.price_SIM2)


 Info_1: The average Equivalent income is equal to : 938.4596

 Info_2: The average implicit utility with reference income in

           the reference situation is equal to : 4.264959

 *** it is the implicit utility before any changes ***

 Info_3: The Current Implicit Utility is equal to: 4.012598

 Info_4: The average implicit utility with income equivalent in

           the reference situation is equal to : 4.012598

 *** it should be equal to implicit utility in the current situation above ***

 Info_5: The average implicit utility with contemporary income in

           the reference situation is equal to : 4.313749
```

The result indicates, as expected, that the welfare of households decreased between the reference period and current period.

8. Conclusion

easi aims at providing a unified framework allowing to estimate and exploit the Exact Affine Stone Index (EASI) demand system of Lewbel and Pendakur (2009) not currently implemented in R. The EASI demand system has several advantages in comparison to AIDS and QUAIDS. Firstly, its numerical implementation is easier due to the linearity in terms of parameters. Secondly, unobserved preferences are taken into account thanks to additive errors which are interpreted as random utility parameters. Finally, easi Engel Curves may have more complicated shapes as suggested by the possible specification of high order polynomials in log real expenditure in the system.

For each estimation, **easi** allows the choice of one version of EASI model with one, two or all of the following interactions : interactions between prices and log real expenditure, interactions between prices and demographic variables and interactions between log real expenditure and demographic variables.

It moreover offers tools for exploitation of the results and simulations.

Among these tools, **easi** provides methods to retrieve more easily the estimates, the residual, the variance matrix, the fitted budget shares and the summary of estimation. It also develops functions to calculate and draw the Engel Curves and functions to calculate elasticities (price elasticities, income elasticities and demographic elasticities).

Furthermore, it enables simulations, namely the assessment of the impact of price changes, income changes and demographics changes on fitted budget shares and elasticities. Similarly, function for calculation of equivalent income is available in **easi**.

Still, extensions and improvements of the software are under way, notably the inclusion of weights.

Research is continuing in this direction.

References

Blundell˜Richard XC, Kristensen D (2007). "Semi-nonparametric IV Estimation of Shape-Invariant Engel Curves." *Econometrica*, **75**.

Bryan BW, Walker MB (1989). "The Random Utility Hypothesis and Inference in Demand Systems." *Econometrica*, **57**.

Daniel M, Richter MK (1981). *Stochastic Rationality and Revealed Stochastic Preference in Preferences, Uncertainty, and Optimality: Essays in Honor of Leonid Hurwicz.* J.S. Chipman, D. McFadden, and M.K. Richter.

Donald BJ, Matzkin RL (1998). "Estimation of Nonparametric Functions in Simultaneous Equations Models, with an Application to Consumer Demand." In *Cowles Fundation Discussion Paper 1175.*

Gorman WM (1981). *Some Engel Curves in Essays in the Theory and Measurement of Consumer Behavior: In Honour of Sir Richard Stone.* Angus Deaton.

Henningsen A (2011). *micEconAids: Demand Analysis with the Almost Ideal Demand System (AIDS).* R package version 0.6-6, URL http://www.r-project.org,http://www.micEcon.org.

Henningsen A, Hamann JD (2011). *Systemfit: Estimating Systems of Simultaneous Equations.* R package version 1.1-10, URL http://www.r-project.org,http://www.systemfit.org.

Lewbel A (2009). "Demand Systems with and without Errors." *The American Economic Review*, **91**.

Lewbel A, Pendakur K (2009). "Tricks with Hicks : The EASI Demand System." *The American Economic Review*, **99**.

Pendakur K (2008). "EASI made Easier." In *EASI made Easier*. URL www.sfu.ca/pendakur/EASImadeEasier.pdf.

Walter B, Blundell R (2004). "Invertibility of Nonparametric Stochastic Demand Functions." In *Birkbeck Working Papers in Economics and Finance 0406*.

Affiliation:

Stéphane Hoareau
CIRPEE
Université Laval
E-mail: sjlhoaro@yahoo.fr

Guy Lacroix
CIRPEE
Université Laval
E-mail: guy.lacroix@ecn.ulaval.ca

Mirella Hoareau
CIRPEE
Université Laval
E-mail: mirella.sinope@yahoo.fr

Luca Tiberti
CIRPEE
Université Laval
E-mail: luca.tiberti@ecn.ulaval.ca