Dipartimento di Statistica
"Giuseppe Parenti"

WORKING PAPER 2008/04

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

Estimating the Parameters of a CES Production Function in a Regional Environmentally Extended CGE Model Framework: a RESAM Only Based GME Approach

Guido Ferrari, Anna Manca

Università degli Studi di Firenze

Estimating the Parameters of a CES Production Function in a Regional Environmentally Extended CGE Model Framework: a RESAM Only Based GME Approach

Guido Ferrari – Anna Manca

Department of Statistics, University of Florence

Viale Morgagni, 59, Firenze, Italia; phone: +39 055 4237221; fax: +39 055 4223560; e-mail: ferrari@ds.unifi.it; Skype ID: morachioli.

Abstract

This paper deals with the problem of estimating the parameters of a Constant Elasticity of Substitution (CES) production function in the framework of a Computable General Equilibrium (CGE) model.

Usually, after specifying the GE model, the computation, consisting of both calibration and parameters estimation, is carried out based on a Social Accounting Matrix (SAM), in some cases supported by additional accounting information, and on information concerning production activity.

Calibrations is performed on the basis of the SAM and, if the case, on the additional accounting information. Estimation of the parameters, and namely of elasticity of substitution, and of income and prices, of the functions that are used both in production and in consumption spheres is performed by making resort to information concerning production, as specified by time-series or cross-section data on enterprises.

A new approach for the above parameters estimation is proposed here based on the first type of macroeconomic information only, by making resort to the Generalized Maximum Entropy (GME) method, which is used for estimating the parameters of a CES production function based on a Regional Environmentally Extended SAM (RESAM).

Keywords: CES function; CGE models; RESAM; GME.

JEL Classification: D_{57}; D_{58}; D_2; C_3.

1 Introduction

In the framework of the Computable General Equilibrium (CGE) models, general equilibrium models standard computation procedure, consisting of both calibration and parameters estimation[1], is carried out, as far as the calibration is concerned, on a database formed by Social Accounting Matrices (SAM), as the core macro-accounting document that represents in terms of values the economic situation one aims at modelling, and purposively on other accounting supporting information, such as satellite accounts.

[1] The term calibration generally indicates the use of procedures that implement the parameterization of economic models (Dawkins, Srinivasan and Whalley, 2001). Accordingly, in calibrating a general equilibrium model, the numerical values of some model parameters, typically the elasticity of substitution in CES functional forms, are obtained exogenously (on the basis of estimates drawn from the literature, or purposively estimated on the basis of time series or cross section databases (Arndt, Robinson and Tarp (2002)), while others, the calibrated parameters, are endogenously determined so as to reproduce the benchmark data as an equilibrium of the model. For the sake of clarity, we prefer here to use the term calibration to denote the endogenously determined parameters only and to leave the term estimation to indicate the exogenously purposively estimated parameters, both in the framework of the more general model computation procedure.

As regards the estimation of the parameters of the production and demand functions used to model the behaviour in production and consumption spheres of activity[2], i.e., efficiency, share, substitution between value added and intermediate consumption parameters, and the related elasticity of substitution, as well as income and prices parameters and elasticity, it is based on time series and cross section data on enterprises.

A significant example of time series utilization for the estimates in both production and consumption spheres is provided by Mansur and Whalley (1981), who, in the calibration of a CGE model for taxation purposes in the USA, have used time series data provided by the US Department of Commerce for estimating a Constant Elasticity of Substitution (CES) production function and a Linear Expenditure System (LES).

Similarly, Jorgenson (1984), has calibrated a CGE model for the USA with production and demand functions estimation based on time series data provided by the US Centre Bureau's National Income and Product Account.

Another worth quoting example of estimates based on time series and cross section data is given by the Centre of Global Trade Analysis (GTAP) of Purdue University, which provides the World Bank (WB) and the Organization for Economic Cooperation and Development (OECD) with production and demand elasticity for the calibration of their standard large CGE model composed by numerous countries.

A variant to the estimation approach is represented by the work by Kehoe, Polo and Sancho (1995), who estimated the elasticity through cross section data in different years, calibrated the related CGE models and carried out a sensitivity analysis to compare the results, as an informal basis for eventually revising the substitution elasticity estimates.

---

[2] Usually Constant Elasticity of Substitution (CES), Cobb-Douglas (CD), Linear Expenditure System (LES), and Almost Ideal Demand System (AIDS) are used.

An interesting approach to estimation is provided by Arndt, Robinson and Tarp (2002), who used time series data joined to prior information on the elasticity to perform a Generalized Maximum Entropy (GME) estimation of trade elasticity associated to a CES aggregated function. These authors utilised a 1991-1996 time series of imports, exports, tariff revenue, total production, marketing margins, intermediate consumption and household consumption for 186 commodities provided by the National Statistical Institute of Mozambique.

Arndt, Liu, and Hertel (2003) followed the same Maximum Entropy (ME) methodology to estimate Armington substitution elasticity in a relatively standard CGE model including 10 countries and focusing on East Asia trade.

In countries, regions or sub-areas where the statistical system is poor and can't provide adequate information, time series or cross section data is not always easy to find, and therefore elasticity cannot be directly statistically estimated and are taken from "external" sources, that is, from sources other than these data.

In some cases, even in countries with well developed statistical systems, the level of detail of the time series or cross section data may not coincide with that of the units classification in the model, which prevents their use for statistical estimation.

This drawback might affect the Environmentally Extended CGE (EECGE) models as well, recently increasingly utilized, as regards the functions through which environmental substitution elasticity has to be estimated.

By and large, the problem is solved through imputation of figures taken from similar contexts: concretely, through a sort of average or central tendency of estimates from a literature survey. This solution is a non-statistical one, as no estimation is performed, being based on analogy instead. Since the context is the general equilibrium, this approach seems reasonable - although being heavily criticized, for example by Jorgenson (1984) - but reveals some drawbacks, the most relevant of which is just represented by the need of seeking the missing elasticity from analogous situations, often difficult to identify.

According to Mansur and Whalley, this is what was done by Pigott and Whalley (1980), who took the substitution elasticity by Caddy (1976), whose paper represents one of most widely used sources in CGE computation, and by Stern, Francis and Shumacher (1976), for the case of an aggregate import and export CES price elasticity. These elasticity are used as point estimates to approximately compute the model at the benchmark equilibrium.

This approach has been followed also by De Melo and Robinson (1981), who have taken trade elasticity from Hickman and Lau (1973) and Alaouze (1977).

Shoven and Whalley (1992) didn't estimate income and price elasticity, and made resort to the above procedure as well.

Mc Kitrick (1998), in his econometric criticism of CGE modelling, pointed out the weakness of the imputation procedure by stressing that, when the elasticity are not estimated and figures taken from other sources are used, the whole CGE model can be seriously affected. Indeed, as stressed by Wing (2004) as well, the empirical foundations of the CGE models become so weak as to question their credibility, so that they are often regarded as black boxes[3].

At our knowledge, no attempt has been made, when time series or cross section data are not available, or even when they are, to use the information contained in the SAM only to estimate the parameters of production or demand functions.

---

[3] Moreover, he claimed that, as shown by Perroni and Rutherford (1995) as regards production, even when the above data is available, flexible forms for the production function should be preferred in order to get estimates with better overall statistical properties.

This paper's aim is to fill this gap in the estimation of the parameters of a CES production function, namely, the efficiency parameter, the share parameter, the substitution parameter between value added and intermediate consumption, and therefore, the elasticity of substitution.

This is done in the framework of the computation of a Regional Environmentally Extended CGE (RECGE) model for Sardinia, a region of Italy, by means of a GME approach based on a Regional Environmentally Extended SAM (RESAM) for Sardinia, 2001 (Ferrari, Garau, Lecca, 2007) only.

This paper's structure is as follows.

In section 2, the GME estimation of the CES production function efficiency, share and substitution parameters, as well as substitution elasticity in a RECGE model framework based on the above RESAM is carried out. In order to adapt the GME approach to the peculiarity of the case of ill-posed database represented by the 2001 Sardinia RESAM, the overall characteristics of the approach itself are first discussed.

In section 3, the description of the database is carried out.

In section 4, the estimation of the above parameters is performed and the analysis of the results is carried out.

In section 5, some concluding remarks are drawn, along with tentative indications of possible further fields of utilization of the methodology.

2 Estimating CES Function Parameters Through the GME Approach in a RECGE Framework Based on a RESAM for Sardinia Only

As said in the Introduction, the estimation of the CES function parameters in a RECGE framework through the GME approach by using the information contained in a RESAM for Sardinia only can be regarded as a specific case of the more general problem of estimating the parameters of a linear model through GME, we are going to discuss below.

2.1 The GME Approach

The estimation of the parameters of a model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, with $\mathbf{y}$ a Nx1 vector, $\mathbf{X}$ a N×K matrix, $\boldsymbol{\beta}$ a K×1 vector and $\mathbf{u}$ a N×1 vector, based on a RESAM in a RECGE model framework is a typical ill-posed problem, as there are not enough degrees of freedom and therefore any traditional econometric solution, such as, a maximum likelihood estimation, would fail. In fact, one needs to estimate β parameters for each of the N branches, that is β_i parameters (i=1,…,N).

An efficient solution is to use GME (Golan, Judge and Miller, 1996; Golan, Judge, and Perloff, 1996; Paris and Howitt, 1998; Golan, Perloff and Shen, 2001; Arndt, Robinson, and Tarp, 2002), which allows to use all the available information without making any parametric assumption on the error term and therefore is robust for a general class of error random variables.

This approach generalizes Jaynes proposal (Jaynes 1957a, 1957b, 1994) of maximizing Shannon entropy measure, H(**p**) = $-\sum_{i=1}^{N} p_i \log(p_i)$, N=number of observations, subject to available data and the probability constraints in order to estimate the probability associated to the unknown parameters.

GME approach (Golan, Judge and Miller, 1996), maximizes the joint entropies (the so-called dual loss objective function) of both signal, represented by the deterministic part of a model and noise, represented by its error distribution, subject to noisy moment representation, that is, to the transformed linear model and to adding-up probability condition.

In other words, let's define:

a discrete random variable $\beta_i = \mathbf{Zp}$, with $\mathbf{Z}$ a N×(N×G) support discrete block diagonal matrix $\begin{bmatrix} z_{i1} & 0 & 0 & . & . & 0 \\ 0 & z_{i2} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & z_{iG} \end{bmatrix}$, i=1,…,N, and with G that represents the dose of a priori information, so that **p** is the N×G-dimensional probability vector associated to the parameter vector β_i through the transformation matrix **Z** (the support space), that is, **p** = F(β_i);

a discrete random variable **u** = **Vw**, with **V** a N×R support discrete block diagonal matrix $\begin{bmatrix} v_{i1} & 0 & 0 & . & . & 0 \\ 0 & v_{i2} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & v_{iR} \end{bmatrix}$, i=1,…,N, and with R that represents the dose of a priori information, the convex hull of the state space of **u**, so that **w** is the probability vector associated to the vector **u** through the transformation matrix **V** (the support space), that is, $\mathbf{w} = F(\mathbf{u}) = F(\mathbf{y}\text{-}\mathbf{X}\beta_i)$.

If $\delta = [\beta_i,\mathbf{u}]$ is a (N×2) joint vector, and $\mathbf{B^*} = \mathbf{ZV}$ a convex set, GME maximizes the dual loss objective function:

$$F(\delta)=F(\beta_i)+F(\mathbf{u})=F(\beta_i)+F(\mathbf{y}\text{-}\mathbf{X}\beta_i)=H(\mathbf{p})+H(\mathbf{w})=-\sum_{i=1}^{N}\sum_{g=1}^{G} p_i \log(p_i)-\sum_{i=1}^{N}\sum_{r=1}^{R} w_i \log(w_i)$$

subject to $\mathbf{y} = \mathbf{XZp} + \mathbf{Vw}$, and normalization: $(\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{G}})\mathbf{p} = (\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{R}})\mathbf{w} = \mathbf{i_N}$.

were H(**p**) and H(**w**) are the entropy measures for the signal and noise respectively, and are defined over the support space **Z** and the joint space **B*** such that $\beta_i = E_{\mathbf{p}}(\beta_i)$ and $\mathbf{u} = E_{\mathbf{w}}(\mathbf{u})$[4].

Besides being capable to exploit all the available information regardless the sample size, the GME approach reaches a unique solution which is assured by the strict convexity of the dual loss objective function and by a positive definite Hessian matrix (Golan, Judge and Miller, 1996).

[4] If β_i, i=1,…,N is the expected value over the support space **Z** and **p** a G-dimensional proper probability distribution defined on the support space **Z**, then $\beta_i = \sum_{d=1}^{G} p_g z_g = E_{\mathbf{p}}(\beta_i)$. Similarly: if **u** is the expected value over the support space **V** and **w** an R-dimensional proper probability distribution associated to the vector **u** through the transformation matrix **V**, then $\mathbf{u} = \sum_{r=1}^{R} w_{kr} v_{kr} = E_{\mathbf{w}}(\mathbf{u})$. In this way, the observed data **y** are viewed as the mean processes β and **u** with a probability distribution *P* that is defined on **Z** and **B***, respectively.

2.2 CES Function Parameters Estimation Based on 2001 Sardinia RESAM

The CES production function whose parameters we are going to estimate based on a Sardinia RESAM only is the following:

$$\log Y_i = \log \alpha_i - \frac{1}{\rho_i} \log\left[\delta_i X_{1,i}^{\ -\rho_i} + (1-\delta_i) X_{2,i}^{\ -\rho_i}\right] + \varepsilon_i; \qquad \text{i=1,...,N} \tag{1}$$

where Y_i is the total output of branch *i*, $X_{1,i}$ the value added, $X_{2,i}$ the intermediate consumption, α_i the efficiency parameters, ρ_i the substitution parameters between value added and intermediate consumption, δ_i the share parameters, and ε_i the error terms. The parameters ρ_i are a transformation of the elasticity of substitution σ_i: $\rho_i = 1/\sigma_i - 1$.

The dual loss objective function to be maximized subject to $\log\mathbf{Y}=\log\mathbf{XZp}+\log\mathbf{Vw}$ and normalization $(\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{S}})\mathbf{p} = (\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{D}})\mathbf{q} = (\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{M}})\mathbf{b} = (\mathbf{I_N}\otimes\mathbf{i}'_{\mathbf{R}})\mathbf{w} = \mathbf{i_N}$, is defined as:

$$\begin{aligned} H(\mathbf{p},\mathbf{q},\mathbf{b},\mathbf{w}) = & -\sum_{i=1}^{N}\sum_{s=1}^{S} p_{i,s}\log p_{i,s} - \sum_{i=1}^{N}\sum_{d=1}^{D} q_{i,d}\log q_{i,d} - \sum_{i=1}^{N}\sum_{m=1}^{M} b_{i,m}\log b_{i,m} - \\ & + \sum_{i=1}^{N}\sum_{r=1}^{R} w_{i,r}\log w_{i,r} \end{aligned} \tag{2}$$

where **p, q, b**, and **w** are the a priori information, i.e., the prior probability distributions on the supports about, respectively, the efficiency parameters α_i, the substitution parameters ρ_i, the share parameters δ_i and the error terms ε_i of the CES production function; *S, D, M* and *R* are the support spaces of, respectively, **p, q, b**, and **w**.

Indeed, the entropy in **p, q, b**, and **w** is expressed by defining each parameter, α_i, ρ_i, δ_i and the error term ε_i, as a (discrete) random variable represented by a convex combination of the related (finite and discrete) support space represented by a discrete support matrix defined by bounding the unknown parameters and disturbances and by using these bounds to specify it, and a weight, represented by a vector of probabilities associated to the support[5]. As no information about the prior probability distributions of the parameters is available, let's assume that they are discrete uniform random variables.

Conversely, a priori information about α_i, ρ_i, δ_i and ε_i can be based on economic theory, evidence from similar analyses, production functions structure and behaviour.

To the purpose of maximizing (2) subject to noisy moment representation of (1), it is necessary first to specify the entropy through re-parameterisation of all the unknown parameters.

Let's focus on the substitution parameters ρ_i first. Its support space can be defined as a set of discrete points:

$$z_i = (z_{i,1}, z_{i,2}, \dots, z_{i,d}, \dots, z_{i,D})' \qquad i=1,\dots N;\ d=1,\dots,D. \tag{3}$$

Associated to the support space we set the vector of unknown weights, defined as:

$$q_i = (q_{i,1}, q_{i,2}, \dots, q_{i,d}, \dots, q_{i,D})' \qquad i=1,\dots N;\ d=1,\dots,D \tag{4}$$

[5] By varying the weights in (2) one can improve the precision term, that is the term which measures the deviations of the estimated parameters from the prior ones, and the prediction term, that is the difference between predicted and observed total output of the variables (the error term).

that must satisfy the probability constrains $\sum_{d=1}^{D} q_{i,d} = 1$ and $q_{i,d} \geq 0$.

Now, we are in a position to write the ρ_i a priori parameters:

$$\rho_i = \sum_{d=1}^{D} z_{i,d} q_{i,d}; (i = 1,..., N) = \begin{bmatrix} \mathbf{z}_{i1} & 0 & 0 & . & . & 0 \\ 0 & \mathbf{z}_{i2} & 0 & . & . & 0 \\ . & . & . & . & . & . \\ 0 & 0 & 0 & . & . & \mathbf{z}_{ND} \end{bmatrix} \begin{bmatrix} \mathbf{q}_{i1} \\ \mathbf{q}_{i2} \\ . \\ \mathbf{q}_{iD} \end{bmatrix} = \mathbf{Zq} \quad (5)$$

where $\mathbf{Z}$ is a N×(N×D) block diagonal matrix for the support space of ρ_i and $\mathbf{q}$ is a (N×D)-dimensional vector of weights.

In this way, each parameter set is converted into a well-behaved set of proper probabilities defined over the support space as a convex combination of points $z_{i,d}$ with weights $q_{i,d}$.

Similarly, the efficiency parameter α_i, the distribution parameter δ_i and the error term ε_i can be re-parameterized following the same procedure as for ρ_i.

In particular: $\alpha_i = \sum_{s=1}^{S} a_{i,s} p_{i,s}$, where $a_{i,s}$ is the support space and $p_{i,s}$ its weight, $\delta_i = \sum_{m=1}^{M} t_{i,m} b_{i,m}$, where $t_{i,m}$ is the support space and $b_{i,m}$ its weight and $\varepsilon_i = \sum_{r=1}^{R} v_{i,r} w_{i,r}$, where $v_{i,r}$ is the support space and $w_{i,r}$ its weight.

By using (5) and the above transformations for α_i, δ_i and ε_i, formula (1) can be re-written as follows:

$$\log Y_i = \log\left(\sum_{s=1}^{S} a_{i,s} p_{i,s}\right) - \frac{1}{\sum_{d=1}^{D} z_{i,d} q_{i,d}} \log\left[\left(\sum_{m=1}^{M} t_{i,m} b_{i,m}\right) X_{1,i}^{-\sum_{d=1}^{D} z_{i,d} q_{i,d}} + \left(1 - \sum_{m=1}^{M} t_{i,m} b_{i,m}\right) X_{2,i}^{-\sum_{d=1}^{D} z_{i,d} q_{i,d}}\right] + \sum_{r=1}^{R} v_{i,r} w_{i,r} \tag{6}$$

This equation represents the consistency constraint, that is the CES production function based on RESAM and on a priori information.

After inserting the probabilities adding-up constraints

$$\mathbf{p}' \cdot \mathbf{p} = \mathbf{q}' \cdot \mathbf{q} = \mathbf{b}' \cdot \mathbf{b} = \mathbf{w}' \cdot \mathbf{w} = 1 \tag{7}$$

maximization of (2) subject to (6) and (7) leads to a unique solution for the weights. Each estimated $\hat{\mathbf{q}}$, $\hat{\mathbf{p}}$, $\hat{\mathbf{b}}$ and $\hat{\mathbf{w}}$ is optimal and represents the posterior probability distribution on the support that satisfies the observations and is the closest to the prior distribution.

By substituting the above optimal solution for $\hat{\mathbf{q}}$, $\hat{\mathbf{p}}$, $\hat{\mathbf{b}}$ and, $\hat{\mathbf{w}}$ the point GME estimates for the parameters can be derived:

$$\hat{\boldsymbol{\rho}} = \mathbf{Z} \cdot \hat{\mathbf{q}} \tag{8}$$

$$\hat{\boldsymbol{\alpha}} = \mathbf{A} \cdot \hat{\mathbf{p}} \tag{9}$$

$$\hat{\boldsymbol{\delta}} = \mathbf{T} \cdot \hat{\mathbf{b}} \tag{10}$$

$$\hat{\boldsymbol{\varepsilon}} = \mathbf{V} \cdot \hat{\mathbf{w}} \tag{11}$$

The GME estimator, which depends on the Lagrange multiplier for the model constrain, has no closed form solution for $\hat{\boldsymbol{\rho}}$, $\hat{\boldsymbol{\alpha}}$, $\hat{\boldsymbol{\delta}}$ and $\hat{\boldsymbol{\varepsilon}}$; therefore, numerical optimization technique has been used to obtain the solution. To implement the model we have used GAMS (General Algebraic Model Solution) software, which utilises a non liner algorithm to solve the model and finds local solution.

3 Data Description

As stressed in the Introduction, this paper's aim is to estimate the parameters of a CES function based on a Regional ESAM only.

Specifically, the database is represented by the Sardinia RESAM for 2001 (Ferrari, Garau and Lecca, 2007), as reported in Table 1. The aggregation of the CES function corresponds to the upper node of the nested tree in the production sphere of the Sardinia CGE.

This RESAM's structure is the classic one suitable for GAMS calibration. The production sphere is detailed into 12 branches, with 11 purely, standard production subjects indicated through abbreviations as reported in Table 1 (with prior acronyms A and C standing for Activities and Consumption), and 1 environmental subject, Waste management (WASTEMAN)[6].

Value added is specified in Income from dependent labour (LABINC), from self-employment (SEINC) and from capital (CAPINC) (rows 27-29), plus social actual and imputed contributions (rows 30-31).

[6] This RESAM is a somewhat rough one, with a branch on waste only, due to poor information on waste situation in Sardinia. It is being improved through a sample survey that is being conducted on the Sardinian enterprises which treat waste management and disposal.

An interesting feature of this RESAM is represented by the subdivision of income allocation to households in 6 income groups, which allows to better understand the primary income allocation process (rows 32 to 37). The other two customary institutional sectors, firms and government (rows 38 and 39) remaining consolidated.

Rows 40-41 record net taxes on production except on VA (row 42), whereas at row 43 saving less investment is recorded. Finally, rows 44 to 47 record environmental investment, tax on export and export to Italy and to the Rest of the World.

The variables in consistency constraint (6) are:

$X_{2,i}$ (i=1,…,12), intermediate consumption, the row vector $\sum_{j=15}^{26} x_{2ji}$ = [579, 152, 1406, 2623, 754, 2364, 912, 2069, 4796, 3108, 3424, 180];

$X_{1,i}$ (i=1,…,12), value added, the row vector $\sum_{j=27}^{31} x_{1ji}$ = [749, 118, 297, 436, 309, 978, 447, 1325, 5545, 3084, 8328, 93];

Y_i (i=1,…,12), total output, the row vector $(X_{2,i} + X_{1,i})$ = [1328, 270, 1703,3059, 1063, 3342, 1359, 3394, 10341, 6192, 11752, 273].

Table 1 – A RESAM for Sardinia, 2001

	A-AGR	A-MINE	A-FOOD	A-TEXT	A-MACHIN	A-OMAN	A-ENERGY	A-CONSTR	A-RETAIL	A-FINANC	A-PUBLIC	A-WASTE
	1	2	3	4	5	6	7	8	9	10	11	12
1												
2												
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
15	175	0	464	5	0	16	1	1	47	5	16	0
16	0	23	2	1023	1	63	197	25	0	1	1	0
17	122	0	355	28	0	3	0	0	367	5	33	0
18	68	16	14	711	20	198	121	55	371	83	466	11
19	4	9	13	23	227	43	25	131	315	70	155	14
20	11	25	107	127	244	1074	16	909	310	150	148	20
21	33	19	41	111	17	197	267	25	242	62	145	1
22	0	1	3	7	4	17	120	414	62	38	208	3
23	84	26	273	276	112	365	46	192	1144	405	355	22
24	81	25	123	170	117	304	72	291	1517	2002	1272	77
25	1	8	8	16	12	33	7	26	408	286	624	16
26	0	0	3	126	0	51	40	0	13	1	1	16
27	208	66	122	159	135	427	149	434	1928	1031	3535	59
28	288	0	97	3	39	179	0	272	1194	225	529	8
29	168	23	20	215	85	196	229	428	1787	1494	2838	13
30	20	29	48	59	46	155	69	156	502	304	1377	12
31	65	0	10	0	4	21	0	35	134	30	49	1
32												
33												
34												
35												
36												
37												
38												
39												
40	7	3	294	653	5	6	110	43	2	373	3	11
41	-103	-134	-41	-2	-15	-196	-21	0	-735	0	1316	0
42	15	0	130	136	68	180	84	195	663	204	30	4
43	258	137	164	131	48	241	324	103	797	623	1247	8
44	0	0	0	0	0	0	0	0	0	0	0	0
45	3	0	14	22	67	18	0	0	0	0	0	0
46												
47												
48	1508	276	2263	3877	1236	3541	1817	3736	11066	7437	11748	324

	C-AGR	C-MINE	C-FOOD	C-TEXT	C-MACHIN	C-OMAN	C-ENERGY	C-CONSTR	C-RETAIL	C-FINANC	C-PUBLIC	C-WASTE	LABINC	SEINC	CAPINC	IMPSOCONTR	SSOCONTR
	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
1	1508																
2		276															
3			2263														
4				3877													
5					1236												
6						3541											
7							1817										
8								3736									
9									11066								
10										7437							
11											11748						
12												324					
15																	
16																	
17																	
18																	
19																	
20																	
21																	
22																	
23																	
24																	
25																	
26																	
27																	
28																	
29																	
30																	
31																	
32													855	0	859	288	0
33													938	337	1112	316	42
34													435	1099	889	147	136
35													3271	535	1822	1102	66
36													1116	230	542	376	28
37													1638	634	715	548	78
38													0	0	890	0	0
39													0	0	667	0	0
40																	
41																	
42																	
43																	
44																	
45																	
46	204	206	1158	781	1434	2586	33	21	280	566	155	3					
47	170	1055	384	654	1744	909	17	14	310	96	92	3					
48	1882	1537	3805	5312	4414	7036	1867	3771	11656	8099	11995	330	8253	2835	7496	2777	350

	H-9300	H-9300-12400	H-12400-15500	H-15500-24800	H-24800-31000	H-31000	FIRMS	GOV	PRODTAX	PRODSUB	VAT	INV	EINV	EXPTAX	Exp-ITALY	Exp-ROW	TOTAL
	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
1																	1508
2																	276
3																	2263
4																	3877
5																	1236
6																	3541
7																	1817
8																	3736
9																	11066
10																	7437
11																	11748
12																	324
15	53	66	57	97	27	43	0	1				1	0		797	10	1882
16	0	0	0	0	0	0	0	0				14	0		167	20	1537
17	280	347	294	507	142	224	0	5				0	0		945	145	3805
18	78	135	110	207	72	78	0	0				0	0		968	1530	5312
19	74	129	141	261	104	139	0	0				1971	23		459	84	4414
20	156	282	275	514	196	312	0	0				301	0		1560	308	7036
21	105	122	118	178	59	83	0	5				0	0		37	0	1867
22	4	4	4	6	3	4	0	0				2845	0		20	2	3771
23	638	948	1303	2283	849	960	0	27				494	0		258	596	11656
24	86	160	153	397	120	210	0	360				352	0		192	22	8099
25	479	523	556	794	360	427	0	6948				103	0		299	54	11995
26	5	10	9	13	8	13	0	19				2	0		0	1	330
27																	8253
28																	2835
29																	7496
30																	2777
31																	350
32	14	0	0	0	0	0	116	756								10	2899
33	0	33	0	0	0	0	267	1737								24	4807
34	0	0	14	0	0	0	117	759								9	3604
35	0	0	0	23	0	0	189	1230								17	8256
36	0	0	0	0	7	0	58	379								5	2742
37	0	0	0	0	0	6	52	336								5	4011
38	82	90	42	315	108	158	179	44								64	1972
39	339	330	35	1232	414	571	885	2719	1510	-2563	1712	7578		125		21	15464
40																	1510
41																	-2563
42																	1712
43	499	1619	490	1403	264	770	47	0					-23		1727	2785	13661
44																	0
45																	125
46																	7429
47	7	9	3	26	9	13	62	139									5712
48	2899	4807	3604	8256	2742	4011	1972	15464	1510	-2563	1712	13661	0	125	7429	5712	

Legend:

AGR = Agriculture, hunting, forestry and fishing; MINE = Mining and quarrying; FOOD = Food, beverage and tobacco products; TEXT = Textiles, Leather, TEXT and Chemicals; MACHIN = Machinery and transport equipment; OMAN = Other manufacturing; ENERGY = Electricity, gas and water supply; CONSTR = Construction; RETAIL = Wholesale and retail trade; hotels and restaurants; transport and communications; FINANC = Financial intermediation; R&D; other business activities; PUBLIC = Public administration; Education; Health; WASTE = Waste management.

4 Estimation Results and Analysis

GME estimates have been obtained by maximizing the joint objective function (2) subject to the CES production function in equation (6), and the adding-up constraints in equation (7).

Support vectors for the parameters ρ_i, α_i, and δ_i have been set up on the basis of the information from the economic theory on the CES production function (Arrow et alii, 1961), and also from other sources, as will be shown shortly.

As far as the block diagonal support matrix **Z** for the substitution parameters between value added and intermediate consumption, ρ_i, is concerned, it has been defined on the basis of what is to be expected for it in the general production framework, that is, from at least constant returns to scale (ρ_i=1) to increasing returns to scale (ρ_i>1) up to reasonably not more than 3-4 (higher returns to scale, even highly desirable, are unlikely to occur in this kind of production processes). Consequently, we have identified the set of discrete points z_i=[1, 1.6, 2.2, 2.8, 3.4], (i=1,…,5), as, according to Golan, Judge and Miller (1996), the choice of five points represents the greatest improvement in precision. Therefore, the support format is:

$$\mathbf{Z} = \begin{bmatrix} 1.0\ 1.6\ 2.2\ 2.8\ 3.4 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & \cdots\cdots & \cdots\cdots & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 1.0\ 1.6\ 2.2\ 2.8\ 3.4 & 0\ 0\ 0\ 0\ 0 & \cdots\cdots & \cdots\cdots & \cdots\cdots & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 1.0\ 1.6\ 2.2\ 2.8\ 3.4 & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & 1.0\ 1.6\ 2.2\ 2.8\ 3.4 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & \cdots\cdots & \cdots\cdots & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 1.0\ 1.6\ 2.2\ 2.8\ 3.4 \end{bmatrix}$$

As regards the efficiency parameters, α_i, it is known from economic theory that it should be positive. In order to get additional information on its likely values, we conducted a sensitivity analysis. According to it, the best a priori distribution seems to be the one ranging [0.5,2.5], so that the support space **A** has been spanned onto a_i=[0.5, 1, 1.5, 2, 2.5].

As for the support **T** for the share parameter δ_i, as the inputs are two, it's perfectly logic to set up it within the interval [0,1], spanning as an uniform distribution t_i = [0.2, 0.4, 0.6, 0.8, 1].

Regarding the error support space **V**, according to literature (Golan, A., Judge, G. and Perloff, J. (1996); Golan A., Perloff J. M. and Shen E. Z.(2001)), we have used a zero symmetric support, v_i=[-1, -0.5, 0.0, 0.5, 1].

In Table 2 the point estimates of the parameters ρ_i, α_i, and δ_i, along with the elasticity of substitution σ_i are presented.

Table 2 – Parameters estimates by branches

	ρ	**α**	**δ**	**σ**
AGR	1,416	1,585	0,594	0,414
MINE	1,896	1,513	0,603	0,345
FOOD	1,544	0,781	0,620	0,393
TEXT	1,524	0,938	0,623	0,396
MACHIN	1,547	1,498	0,615	0,393
OMAN	1,302	1,679	0,620	0,434
ENERGY	1,480	1,477	0,615	0,403
CONSTR	1,256	1,946	0,611	0,443
RETAIL	1,076	1,741	0,592	0,482
FINANC	1,145	1,633	0,600	0,466
PUBLIC	1,080	1,730	0,555	0,481
WASTE	1,872	1,500	0,605	0,348

First of all, there is evidence that the most efficient branch is Construction, with an α estimated parameter equal to 1.946, firmly shifted toward the upper bound 2.5 of the a priori information set of values. Nearly equally efficient is the Wholesale and retail branch, the branch of services, with a lower efficiency estimate (1.741). Similarly, as for Other manufacturing and Financial branches, with respectively, 1.679 and 1.633.

Thus, for the first two branches, the starting level of production is more than satisfactory, as compared to the expectations, whereas for the second two, the level is a bit above the average expected level of starting production.

Conversely, the Food branch exhibits an efficiency parameter of 0.781, very closed to the lowest bound of the expected efficiency. Somewhat low starting level of production is shown by the Textile, leather, coke, and chemical branch too, with its 0.938 parameter estimate.

As far as the remaining branches are concerned, they reach on average the expected efficiency, i.e., the expected starting level of production, with Waste management taking exactly the median position (1.500).

Let's analyse now the elasticity of substitution between value added and intermediate consumption evidence, starting first from a look at the substitution parameter ρ by branches.

By and large, it takes quite low values as compared to the a priori distribution, which ranges from 1 to 3.4, with the highest values taken by Mining (1.896) and Waste management (1.872).

Substitution parameters estimates higher than unity for all branches stress the concordance between intermediate consumption and value added and total production in all of them, which is a symptom of a certain rigidity in substitution between the two production factors.

Across branches analysis of the elasticity of substitution σ confirms the above, by witnessing a uniformly somewhat low rate of substitution for all branches, with, of course, the lowest ones for Mining and Waste management. The latter being a finding that might have been expected due to the relatively more labour and capital intensive characteristics of the two branches. Similarly, as for Textile and Machinery branches, equally two somewhat capital and labour use oriented branches.

The relative importance of the two inputs, that's, value added and intermediate consumption, in production activity in Sardinia is provided by the estimates of the δ parameter. There is strong evidence of the prevalence of value added over intermediate consumption in all branches – no parameter estimate witnesses of an use of value added lower than 55% out of total inputs use - with Textile and Other manufacturing being the most capital oriented branches, with their 62% of value added use, closely followed by Machinery, Energy and Construction, with 61%. Waste management is a very capital oriented branch too, by totalling a 60% use of it.

In order to measure the information content in the system and the consequent importance of the contribution of data in reducing uncertainty, we have utilised the information index $R^* = 1 - S(\hat{w})$, where $S(\hat{w}) = \frac{-\sum_{\omega} \hat{w}_{\omega} \log \hat{w}_{\omega}}{\Omega \log \Omega}$ is the normalised entropy proposed by Golan, Judge and Miller (1996). S($\hat{w}$) represents the proportion of the remaining uncertainty, as the numerator indicates the entropy related to the data information, while the denominator indicates the maximum level of uncertainty, that is the entropy level of the uniform distribution with Ω outcomes. S($\hat{w}$)∈[0,1] and of course, R*∈[0,1] too. R*=1 means perfect prediction, i.e., that the sampling evidence, that's, the information contained in the database, provides a full contribution to the a priori information to get likelihood, by perfectly combining with it and therefore to obtain exact prediction, whereas R*=0 means complete uncertainty.

In table 3, the information index R* calculated for each parameter is presented.

Table 3 – Information index R*

	ρ	α	δ
	$R^* = 1 - S(\hat{q})$	$R^* = 1 - S(\hat{p})$	$R^* = 1 - S(\hat{b})$
AGR	0,896	0,858	0,884
MINE	0,920	0,949	0,930
FOOD	0,890	0,780	0,878
TEXT	0,882	0,735	0,874
MACHIN	0,899	0,876	0,846
OMAN	0,884	0,900	0,870
ENERGY	0,893	0,880	0,895
CONSTR	0,883	0,919	0,883
RETAIL	0,867	0,975	0,879
FINANC	0,873	0,982	0,875
PUBLIC	0,867	0,969	0,899
WASTE	0,918	0,829	0,927

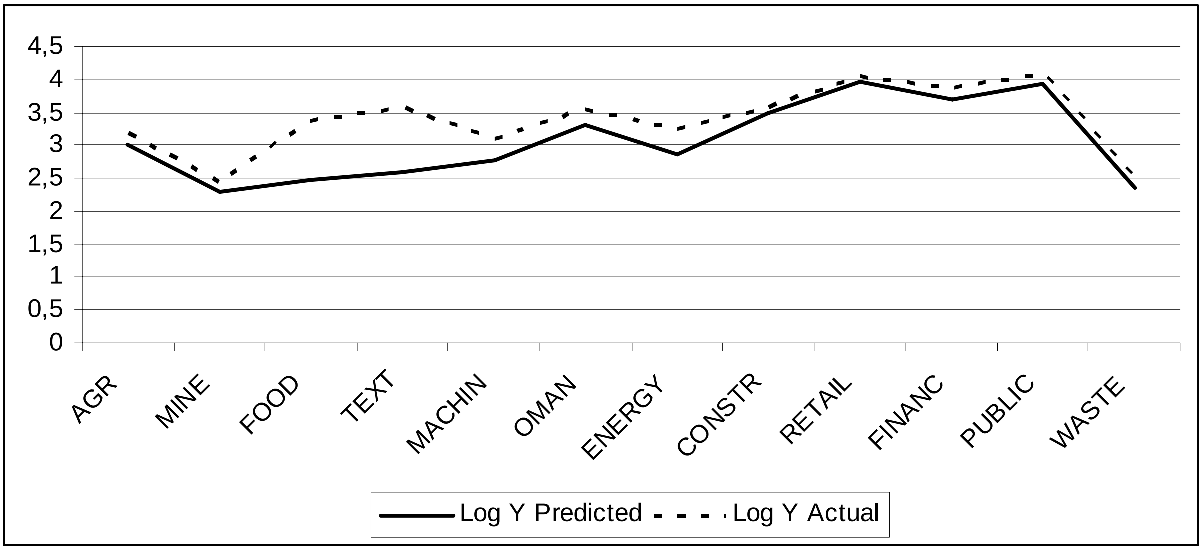
It clearly appears that the information from database greatly contributes to reduce uncertainty, and therefore, to get good output predictions in all branches, as is confirmed by Figure 1, where observed (dotted line) and predicted outputs are shown for each of them.

Indeed, the remaining uncertainty is quite low, accounting for less than 20% for all the three parameters in all branches, except the case for α in Food and Energy branches, where there is evidence of more than 20% remaining uncertainty (respectively, 22% and 26.5%).

On the other hand, as above underlined, these two branches are those which exhibit the lowest level of efficiency.

This explains the bigger differences between observed and predicted outputs as shown in Figure 1.

Figure 1: Total output



5 Concluding Remarks

The approach we propose in this paper is information parsimonious, as it avoids making resort to data sources other than those contained in the RESAM. In this sense, it represents a pure case of ill-posed estimation.

In this situation, the GME estimates of the CES production function parameters globally are satisfactory. Indeed, the information index R* shows that the information from database decidedly contributes to reduces uncertainty. This is a clear response that the road undertaken to base estimation on RESAM only, a priori logically defensible indeed, is well chosen and rich of promises for the future, with the interesting developments it catches sight of as well.

From an economic point of view, the parameter estimates conform to the expectation and turn out to be consistent with the Sardinian economic system.

Efficiency seems to comply the production features of the island, on average. Indeed, efficiency parameter higher than one seem to adequately reflect the technological structure of the production system, characterized by small-medium enterprises, with a not negligible degree of technical flexibility. Waste branch provides a paradigmatic confirmation of this perception.

There is evidence of the prevailing relevance of value added over intermediate consumption in all branches, as was logic to expect, because of the capital intensive oriented type of activity in the island.

Moreover, the two production factors are hardly replaceable each other, as witnessed by the uniformly low rate of substitution between them.

It seems quite redundant to stress that the analysis carried out in this paper is still in progress, as regards both methodology and application.

In particular, many questions related to the GME approach in ill-posed situations such as the present one need to be duly refined and better specified.

First, the thorn of the a priori distribution definition. It is a crucial and still somewhat unsolved problem or, perhaps, a not satisfactorily placed one. Indeed, our opinion is that more can be done; in particular, a deeper analysis of the theoretical assumptions and of their implications is in order. To make resort to literature and or even to a broad economic theory framework doesn't seem enough and statistically based supporting analyses should be submitted to the debate.

Secondly, the problems of the database, basically the possibility to have a good RESAM at disposal, that in our approach is crucial as the matrix represents the sole source of information, is open and claims deeper and more detailed attention.

References

Arndt, C., Robinson, S. and Tarp, F. (2002), Parameter Estimation for a Computable General Equilibrium Model: a Maximum Entropy Approach, Economic Modelling, 19(3), 375–398.

Arrow, K, Chenery, H, Minhas, B and Solow, R (1961), Capital-Labor Substitution and Economic Efficiency, The Review of the Economics and Statistics, 43(3), 225-250.

Caddy, V. (1976), Empirical Estimation of Elasticity of Substitution: a Review, mimeo, Industry Assistance Commission, Melbourne, Australia.

Dawkins, C., Srinivasan, T.N., and Whalley, J. (2001), Calibration, Chapter 58, Handbook of Econometrics, Volume 5, Edited by Heckman J.J. and Leamer E. (Elsevier Science B.V).

De Melo, J. and Robinson, S. (1981), Trade Policy and Resource Allocation in Presence of Product Differentiation, The Review of Economics and Statistics, 63(2), 169–177.

Ferrari, G., Garau, G., and Lecca, P. (2007), An ESAM for Sardinia, mimeo, Dipartimento di Statistica, Università di Firenze.

Golan, A., Judge, G and Miller, D (1996), Maximum Entropy Econometrics: Robust Estimation with Limited Data (Chichester, Wiley).

Golan, A., Judge, G. and Perloff, J. (1996), A Maximum Entropy Approach to Recovering Information from Multinomial Response Data, Journal of the American Statistical Association, 91(434), 841-853.

Golan, A., and Gzyl, H (2002), A Generalized Maxentropic Inversion Procedure for Noisy Data, Applied Mathematics and Computation, 127(2-3), 249-260.

Golan A., Perloff J. M. and Shen E. Z.(2001), Estimating a Demand System with Non-negativity Constraints: Mexican Meat Demand, The Review of Economics and Statistics, 83(3), 541-550.

Jaynes, E.T.(1957a), Information Theory and Statistical Mechanics I, Physics Review, 106, 620-630.

Jaynes, E.T. (1957b), Information Theory and Statistical Mechanics II, Physics Review, 108, 171-190.

Jaynes, E.T. (1994), Probability Theory: the Logic of Science (Cambridge, Cambridge University Press).

Jorgenson, D. (1984), Econometric Methods for Applied General Equilibrium Analysis, in Scarf, Herbert E. and Shoven, J. B. (eds.) Applied General Equilibrium Analysis (New York, Cambridge University Press).

Kapur, J.N. and H.K. Kesavan (1992), Entropy Optimization Principles with Applications (Boston, Academic Press).

Kehoe, T.J., C. Polo, and F. Sancho (1995), An Evaluation of the Performance of an Applied General Equilibrium Model of the Spanish Economy, Economic Theory, 6, 115-141.

Liu, J., Arndt, C. and Hertel, T.W. (2003), Parameter Estimation and Measures of Fit in A Global, General Equilibrium Model, Conference Proceedings–Fourth Annual Conference on Global Economic Analysis I.

McFadden, D. (1963), Constant Elasticity of Substitution Production Functions, Review of Economic Studies, 30, 73-83.

McKitrick, R.R. (1998), The Econometric Critique of Computable General Equilibrium Modelling: the Role of Parameter Estimation, Economic Modelling, 15, 543-573.

Paris, Q. and Howitt, R. (1998), An Analysis of Ill-posed Production Problems Using Maximum Entropy, American Journal of Agriculture and Economic, 80, 124–138.

Perroni, C. and Rutherford, T. (1998), A Comparison of the Performance of Flexible Functional Forms for Use in Applied General Equilibrium Modelling, Computational Economics 11, 245–263.

Preckel, P. (2001), Least Squares and Entropy: A Penalty Function Perspective, American Journal of Agricultural Economics, 83(2), 366–377.

Shannon, C.E. (1948), A Mathematical Theory of Communication, Bell System Technical Journal, 27, 379-423.

Shoven, J. and Whalley, J. (1984), Applied General Equilibrium Model of Taxation and International Trade: an Introduction and Survey, Journal of Economic Literature, 22(3), 1007-10051.

Stern, M., Francis, J. and Schumaker, B. (1976), Price Elasticities in International Trade: an Annotated Bibliography (McMillan, London).

Uzawa, H. (1962), Production Functions with Constant Elasticity of Substitution, Review of Economic Studies, 29, 291-99.

Wing, S. (2004), Computable General Equilibrium Models and Their Use in Economy-wide Policy Analysis, Technical report, Massachusetts Institute of Technology: Joint Program on the Science and Policy of Global Change.