Università degli Studi di Firenze
Dipartimento di Statistica "G. Parenti"
Scuola di Dottorato di Ricerca in Statistica Applicata
XXI ciclo

Entropy Approaches for the Production and Demand Function Parameter Estimation in a Regional CGE Model Framework

Luca Secondi

Tutor	Prof. Guido Ferrari
Co-Tutor	Prof. Amos Golan
Co-Tutor	Prof.ssa Tiziana Laureti
Director	Prof. Guido Ferrari

To my parents

Acknowledgements

As this important phase of my education and research is drawing to the end I would like to thank all those who have assisted and believed in me.
I sincerely thank Prof. Guido Ferrari who supervised me wisely and patiently and gave me support and advice, making this research possible. In particular, I thank him for having taught me so much about science as well as life.
A hearty thanks to Prof. Amos Golan who gave me the opportunity of spending some time at the American University in Washington DC where I was able to take advantage of an intense and profitable period of work. I thank Prof. Golan for his kindness and his precious contribution.
A special thanks to Prof. Tiziana Laureti who was the first to encourage me to take up my doctorate studies. I especially thank her for assisting me with perseverance, professionalism and for help and advice in hard times.
I would also like to thank all the Professors in my Department for the important things they taught me during these years of doctorate study. A special thanks to Prof. Silvana Salvini.
Finally I thank my doctorate workmates with whom I have lived and shared a pleasant and important three years. A special thank you to Elena and Laura.
Last but not least, I am very grateful to my parents, my brother and Alessandra for being close to me during this time. I especially thank my parents for supporting me while I was concluding this important phase of my life. Thank you Alessandra for being so close to me and demonstrating such affection, respect and comprehension.

CONTENTS

Introduction

In the computation of General Equilibrium (GE) models, one of the most debated issues concerns the determination and, consequently, the validity of the parameters describing the behaviour of economic agents, paying particular attention to production and demand functions (Mansur and Whalley, 1984; Shoven and Whalley, 1984, 1992; McKitrick, 1998; Arndt, Robinson and Tarp, 2002; Boys and Florax, 2007; Partridge and Rickman, 2008).
This topic is relevant because there are actually two important phases when computing a GE model: calibration and estimation.
As several authors have pointed out, see as example Hoover (1995), Hansen and Heckman (1996), Browning, Hansen and Heckman (1999), Balistreri and Hillberry (2005), there is still a heated debate in literature concerning the definition and the specification of a "*boundary line*" between calibration and estimation. On this issue, Dawkins, Srinivasan and Whalley (2001) made an interesting proposal in their review article because, according to them, calibration and estimation are closely related and actually in less conflict than one might suppose, as they stress that "*calibration is estimation and estimation is calibration"*. The same authors define calibration as *"... the setting of the numerical values of model parameters relative to the criterion of an ability to replicate a base case dataset as a model solution* ". On the other hand, they refer to estimation as "*… the use of a goodness of fit criterion in the selection of numerical values of model parameters*".
Calibration and estimation in a Computable General Equilibrium Model (CGE) are generally developed from databases represented by Social Accounting Matrices (SAMs), some additional macroeconomic information and other data coming from external sources, such as information resulting from similar contexts or time-series and cross-section data. However, SAM represents the core macro-accounting document, since it sums up the economic situation of the system in question in terms of values.
Specifically, calibration refers to the use of information contained in a SAM in order to express macroeconomic equalities, identities and sometimes behavioural function parameters of the model in value terms. This phase of the computation process does not present any particular difficulty except for the usual problems related to the possibility of obtaining a "good" SAM and of using this information correctly.
This research is mainly focused on the phase of estimation which concerns the estimation of behavioural parameters. As stressed by Devarajan and Robinson (2002) this phase concerns those parameters, like elasticity parameters, related to production, consumption, import demand and export supply functions, which cannot usually be obtained from a single SAM, but for which additional external data or external information is required.

Starting from the earlier applications - which were developed in the 1960s – and up to the 1980s, by using the "*imputation*" procedure previously estimated parameters obtained from similar situations were frequently used in order to obtain the value of the behavioural parameters (Piggott and Whalley, 1980; De Melo and Robinson, 1981; Dixon, Parmenter, Sutton and Vincent, 1982; Shoven and Whalley, 1982). However, "borrowing" the values from existing literature is not a real statistical approach, as no estimation has been performed.

In the most recent years, the estimation has usually been carried out by following the *econometric approach* which refers to above-mentioned use of cross-section and time-series databases in addition to the SAM (Jorgenson, 1984; Jorgenson and Slesnick, 1997; McKitrick, 1998; Adkins, Rickman and Hameed, 2003).

However, both procedures have been widely criticised, and therefore obtaining appropriate values for behavioural parameters is still a debatable problem. In fact, the validity of these parameters, as stressed by Iqbal and Siddiqui (2001) Arndt, Robinson and Tarp (2002), McDaniel and Balistreri (2002) Nganou (2004) and other authors, is a very crucial point since CGE results have shown to be quite sensitive to their values.

The *imputation procedure* has been widely criticised on several grounds by Jorgenson (1984) and McKitrick (1998). On this issue, Wing (2004) stated that, although CGE models are widely diffused and generally accepted, they are still doubted by economists and politicians who have labelled them "*black boxes*", since the imposition of the use of values borrowed from existing literature may not be suitable to the economic context in question. Consequently, the results of the model cannot be meaningfully traced to any particular features of their database or input parameters, algebraic structure, or methods of solution. Similarly, Dawkins et al. (2001), in referring to the values of elasticity, underline that borrowing the values of these parameters from already existing studies represents one of the clearest "*Achille's heels*" of CGE models.

Whalley (1985) and Nganou (2004) pointed out that "borrowed values" for behavioural parameters are unlikely to be appropriate for the country the CGE model is trying to represent and, therefore, they do not usually suit the specific context of study.

Partridge and Rickman (1998; 2008) outlined the same problem for the regional CGE (RCGE) models, stressing that the same external parameter values found in national or international contexts are routinely used also in the regional context.

As regards the *econometric approach,* Jorgenson (1984), in his review concerning the procedure of imputation in a CGE model, was the first author to propose a model characterized by a computation process in which the behavioural parameters regarding the fields of production, consumption and commerce were estimated by using annual time-series data in an attempt to solve the problems related to the imputation procedure.

Jorgenson's study has opened a "new" path for CGE models characterized by the development of econometric methods for estimating behavioural parameters necessary for the equilibrium of the model. The same approach has been reviewed, further developed and improved by, Jorgensen, Slenick and Wilcoxen

(1992), Kehoe, Polo and Sancho (1995), Arndt, Robinson and Tarp, (2002), Boys and Florax (2007) and other authors.
However, this approach has frequently been criticized as well. As stated by Blackorby and Russel (1989), the estimation of specific input parameters, such as elasticity, is affected by conditions and assumptions specific to each estimation process and therefore not easily and immediately extensible. Moreover, by using external data, some problems of classification and comparability of data can arise.

The main objective of this research is to overcome the weaknesses of these two procedures for the determination of behavioural parameters concerning production and demand functions, by using the information contained in a SAM only. In particular, this will be done in the framework of RCGE models.
To be precise, our idea is to unify the calibration and the estimation phases into a simultaneous estimation process, so that the computation process is completed using the regional SAM (RSAM) only. In this way, we will also achieve a twofold additional result. On one hand, we will emphasize the important role that calibration plays in carrying out a "self-contained" computation process. On the other hand, we will enhance the existing debate in literature by proving that the calibration and the estimation phases are much more connected than they have been considered up to now. In such a way, our approach confirms the statements of Dawkins et al. (2001) and Balistreri and Hillberry (2005) who underlined that "calibrators and estimators actually communicate much less than they should and that this lack of communication leads to a lack of research".
The sole use of this type of macroeconomic information for the parameter estimates is equivalent to a "*limited information*" approach and specifically to an "*ill-posed*" situation since a RSAM does not contain enough information for a significant statistical estimation.
Ill-posed situations have been dealt with in literature by taking advantage of the class of Information-Theoretic methods. All these methods, as underlined by Golan (2008), pursue the "classical" objective of extracting all of the available information from the data with minimal assumptions on the underlying distribution generating the data.
In this research we specifically refer to the Generalized Maximum Entropy (GME) and Generalized Cross Entropy (GCE) estimators proposed by Golan, Judge and Miller (1996a), which are based on the entropy criterion introduced by Shannon (1948) and its extensions, such as the Kullback-Leibler's divergence measure (1951), and represent a generalization of the Maximum Entropy (ME) principle introduced by Jaynes (1957a, 1957b).
Golan et al. (1996a) stated that these methodological approaches allow us to solve the problem of an econometric model that otherwise would be ill-posed (Paris and Howitt, 1998) and to obtain estimates that are consistent with "*what we know*", that is the original dataset.
GME uses all and only the available data and it lets the data speak for itself without imposing unnecessary or arbitrary assumptions or restrictions and does not require any restrictive moment or distributional error assumptions. The starting point of GME's optimization problem is the specification of a finite and

discrete support for each of the unknown parameters and for the error component. By taking into account in the data constraint the disturbance terms, the GME can be considered as an extension of the classical ME approach.
The GME estimates are obtained as the solution of a maximization problem subject to the available data constraints. In fact, due to this optimization procedure a mathematical problem is transformed into a deductive problem. (Golan et al., 1996a). In other words, the GME method leads to the maximization of the joint entropy of both the signal, represented by the deterministic part of the model (the data), and the noise which represents the stochastic component.
The solution of the optimization problem results in the best prediction of unknown parameters by transforming the empirical evidence into a probability distribution, which describes the state of our knowledge. By proceeding in this way, the entropy can be viewed as a measure of uncertainty associated with the probability distribution which is consistent with the available information.
Additional information concerning the unknown parameters may be expressed in the form of prior probability distributions on the supports identified for the unknown parameters. The introduction of "a priori" information in the form of a probability distribution leads to the GCE estimator and specifically to a minimization problem. Actually, the entropy objective is used here to find the set of "posterior" distributions on the supports that satisfy the observations and are "closest" to the prior distributions.
Since their introduction, a notable number of applications of GME and GCE have appeared in literature referring to various types of economic contexts and situations.
Concerning production and demand functions, interesting applications are provided by Lence and Miller (1998), Paris and Howitt (1998) , Fraser (2000), Balcombe, Rapsomanikis and Klonaris (2003), Golan, Perloff and Shen (2001), Howitt and Msangi (2006).
Moreover, whenever a CGE framework is called for, GME and GCE estimators appear to be among the most appealing tools. On this issue, interesting applications are provided in the studies by Arndt, Robinson and Tarp (2002), Nganou (2004), Ferrari and Manca (2008).
Regarding the empirical application of this research, the benchmark dataset under which the estimates will be carried out is represented by the RSAM of Sardinia for the year 2001 (Ferrari, Garau and Lecca, 2007).
To put the analysis in the regional context, where additional data is much less readily available than at national level, means facing further complications due to peculiarities of a region such as the degree of specificity or mobility of the factors and the regional product differentiation (Partridge and Rickman, 1998). The greater degree of openness of a regional economy which not only trades with other countries but also with the other regions in the same country, is an example of this aspect. At the same time, labour mobility seems to be more accentuated among regions rather than among countries.
Estimation procedures used for obtaining production and demand function parameter values are strictly related to the behavioural sphere in question.

Concerning production, we will introduce a *three-step estimation process* based on the entropy estimation approaches in which additional information is gradually incorporated into the estimation process, because it is necessary to combine the available RSAM data, the calibrated RSAM information and the economic theory requirements.
By following this strategy, in the *first step* no "a priori" information regarding the order of sizes and magnitude of the parameter values will be introduced, according to the classical GME approach.
In the *second step* additional information will be introduced, by using the results obtained in the previous step together with the RSAM information and the economic theory.
Thirdly, *the final step* is based on GCE philosophy where prior information will be introduced both in terms of parameter support bounds, as a result of the second step, and in terms of an estimated prior probability distribution obtained by the ME estimator.
Concerning demand, we will take advantage of the classical GME estimation method which does not assume any "a priori" information. In this context, we will explain how we estimate systems of demand equations for the different household groups portrayed in the Sardinian RSAM.
As mentioned above, the most important aspect of this thesis is the obtainment of the estimates of behavioural parameters, which will be totally suitable for the computation process of the RCGE model for Sardinia, since the parameter values will be obtained in a "self-contained" approach where no external information is assumed.
To our knowledge, there has been no research for the estimation of behavioural parameters using this approach and exclusively the data contained in the SAM.

The remainder of this work is structured as follows.

The first chapter deals with CGE models by analyzing the "state of the art", the scientific debate concerning them and the different approaches used in literature to determine behavioural parameter values for both production and demand functions. Taking into account the faults and defects that these approaches have proved to have, we will illustrate the motivation and the main contribution of our approach.

The second chapter deals with the class of Information Theoretic estimators by focusing on the entropy based methods. Beginning with Shannon's entropy criterion and the ME principle developed by Jaynes, a functional description of the GME and the GCE estimators is provided, together with a description of the most important and interesting applications in the context of production and demand as well as in the CGE and RCGE framework.

The third chapter describes the main features of our proposed approach in detail. After illustrating the key roles that production and demand functions have in the RCGE framework, we will describe the chosen functional forms for the Sardinian

context. This chapter also includes an explanation of the main characteristics of the Constant Elasticity of Substitution (CES) and the Working-Leser functions, which are the selected functional forms for the production and demand spheres, respectively.
Furthermore, we will make a detailed explanation on the economic meaning of the SAM as well as on the specific structure of the RSAM of Sardinia for the year 2001, which represents the database used for the estimates.

In the fourth chapter the estimation of the CES production function, between value added and intermediate consumption, will be performed. This chapter also deals with the proposed *three-step* estimation procedure by providing the most essential aspects and some indications about future developments and further improvements. Finally, the parameter estimates, together with the obtained elasticity of substitution will be discussed.

In the fifth chapter, we will carry out the Working-Leser demand system estimation. This chapter deals with the classical GME estimator approach by illustrating how this estimator works and discussing the obtained expenditure elasticity, suitable for feeding the RCGE model equations.

In the concluding remarks, we will point out the main results obtained and how our approach can be included in the existing heated debate in literature. Finally, we will make further indications concerning some plausible improvements as well as other fields of application of the proposed estimation strategy.

Chapter 1

The Economic Framework: CGE Models and the Computation Process

"…In many areas elasticity estimates differ in both size and sign, while for a number of the issues
in which applied modellers are interested in, no relevant elasticity estimates exist.
The choice of elasticity values in applied models is therefore frequently
based on contradictory, or little or no empirical evidence.
This obviously undetermines confidence in model results. [Whalley, 1985]

1.1 CGE Models: the "State of the Art" and the Scientific Debate

CGE models, have been widely used for quantitative analysis of global economic issues and are still among the most useful tools in applied economics for simulating alternative economic policies.

CGE models actually represent the most important instrument for combining the theoretical/abstract general equilibrium structure formalized in the 1950s by Arrow and Debreu and realistic economic data in order to convert an abstract representation of an economy into realistic models of actual economies (Shoven and Whalley, 1984). At the same time, numerically speaking CGE models enable us to solve for the levels of supply, demand and price that support equilibrium across a specified set of markets (Wing, 2004).

As underlined by Wise, Cooke and Holland (2002), CGE models explicitly capture the behaviour of the various agents (households, firms, government, and rest of the world), the institutional framework (fiscal system and transfer mechanisms), and the market clearing processes (prices and quantities). They provide an internally consistent representation of the economic structure through the specification of a system of simultaneous equations following the Walrasian general equilibrium system.

The first empirically based CGE model was implemented by the economist Leif Johansen in 1960, who formulated a multi-sector, price-endogenous model, for Norway, analyzing resource allocation issues[1]. Johansen retained the fixed coefficient assumption in modelling demands for intermediate goods, but he utilized the Cobb-Douglas (CD) production functions for modelling the substitution between capital and labour services and technical changes (Jorgenson, 1983).

---

[1] Dixon (2006) underlines that on a broader definition, CGE modelling starts with Leontief's input-output models of the 1930s and includes the economy-wide mathematical programming models of Sandee.

Initially confined to universities and research institutions, today CGE models are usually used by governments in policy formulation and debates (Devarajan and Robinson, 2002).

As stressed by Bergman (1982), following Johansen's study, the development and application of CGE models has become a rapidly growing field in economics.

Much research has been carried out since the 1970s focusing on empirical applications for the analysis of trade, taxation, income distribution, structural adjustment, industrial policies, environmental issues and so on, both in developed and developing countries.

Various types of CGE models can be developed. The choice depends on the particular objectives of the study for which the model will be used. On this issue, Devarajan and Robinson (2002) argued that models destined for use in policy analysis should meet a number of the following desirable features: i) policy relevance; ii) transparency in the links between policy variables and outcomes; iii) timeliness; iv) validation and estimation relevant to the policy issue, that is to mean the model must be determined to achieve accurate results for the domain of potential policy choices; v) diversity of approaches[2].

Several reviews concerning the different application of CGE models exist. Shoven et al. (1984) reviewed the early national CGE literature highlighting the main features of applied general-equilibrium tax models. In their paper, they stressed that earlier analytic work with these models examined the distortional effect of taxes, tariffs and other policies, along with functional incidence issues. On the other hand, more recent applied models cited in their review provide numerical estimates of efficiency and distributional effects.

De Melo (1988) analyzes the application of CGE models for the quantification of trade policy scenarios. Decaluwè and Martens (1988) and Bandara (1991), examine applications of CGE models for developing economies. Bhattacharyya (1996) reviewed the implementation of CGE models concerning energetic and environmental problems.

Concerning RCGE models, our field of interest, an interesting review is provided by Partridge and Rickman (1998) in which they also argued that practical approaches should be taken to overcome data limitations in the computation of RCGE models.

André, Cardenete and Velaszquez (2005), by using a RCGE model, analyzed an environmental tax reform and the double dividend hypothesis for CO_2 and SO_2 policies for the region of Andalusia in Spain.

Saveyn and Regemorter (2007) developed and applied a RCGE model for environmental and energy policies for the federal structure of Belgium. His RCGE model differs from the national CGE models as they take into account the interregional mobility of labour, the common product market across the regions and the explicit modelling of two government levels within one Nation.

[2] In this respect, the same Authors stressed that the experience of the past twenty years about applied CGE models seems to suggest that "it is better to have a good structural model capturing the relevant behaviour of economic actors and their links across markets, even if the parameters are imperfectly estimated, because the domain of applicability of such models makes them more useful for policy analysis than other stylized models".

Partridge and Rickman (2008) discussed the likely reason for the limited use of RCGE models for economic analysis, particularly for small region. They also proposed methodological improvements in order to lead to a wider use of RCGE models in economic development practice.
In actual fact, analyzing a sub-national context within the RCGE framework means facing additional issues which arise with the appropriate definition of a region such as the degree of factor specificity/mobility and regional product differentiation (Partridge and Rickman, 1998).

Despite their extensive use and general acceptance, RCGE models (and equivalently national CGE models) have been frequently criticized for the weak empirical foundation on which they are based[3] (Jorgenson, 1984; Singleton, 1988; Hoover, 1995; McKitrick, 1998; Iqbal and Siddiqui, 2001; Balistreri, McDaniel and Wong, 2002; Boys and Florax, 2007; Hertel, Hummels, Ivanic and Keeney, 2007).
One of the most debated issues concerns the selection and, consequently, the validity of the key behavioural parameter values used in the computation process since as stated by Wigle (1986) and other authors, the selection of parameter values may be highly subjective and therefore raises natural scepticism regarding the reliability of the resulting simulations.
On this issue, Wing (2004) by stressing the wide diffusion of CGE models underlines that these models are still viewed with suspicion in economics and political communities and they are defined as "*black boxes*" since the results cannot be meaningfully traced to any particular features of their data base or input parameters, algebraic structures, or methods of solution.
The importance of this problem has been recognized by several authors and various alternative approaches have been proposed for assessing the simulation uncertainty induced by parameter uncertainty (Pagan and Shannon, 1985; Wigle, 1991).
For all these reasons, the behavioural parameter values play a crucial role in the functioning of the model and consequently affect the results of policy and external shock simulations since their values, as stressed by Annabi, Cockburn and Decaluwè (2006) critically determine the magnitude of response to different exogenous shocks.
Shoven and Whalley (1984) were the first authors to point out the critical role of parameter selection underlying the associated difficulties in the simulation results. These authors stressed that up to the 1980s the most widely used procedure for assessing the reliability of the simulations consisted in performing a few alternative simulations with different parameter values and that the most common procedure was to choose a central case specification, around which a sensitivity analysis could be performed.
The behavioural parameter selection methods applied can be summarized into two main approaches. In short, the first approach known as "*imputation procedure*"

[3] Schmalensee, Stocker and Judson (1998) underline that this problem is not really confined to CGE models, but it has been recognized for complex simulation models in general.

refers to taking the behavioural parameter values from existing studies and to insert them into the CGE model in question. The second approach, known as the "*econometric approach*", refers to the determination of the behavioural parameter values through an estimation procedure based on "external source" data. However, both of the above mentioned procedures have been widely criticised, and therefore obtaining appropriate values for behavioural parameters is a debatable problem even today. In fact, the estimation of these parameters, as stressed by Arndt et al. (2002), McDaniel and Balistreri (2002), Nganou (2004) and other authors, is a very crucial point since CGE results have been shown to be quite sensitive to the value of these parameters. In the next paragraphs, an explanation of the two approaches can be seen.

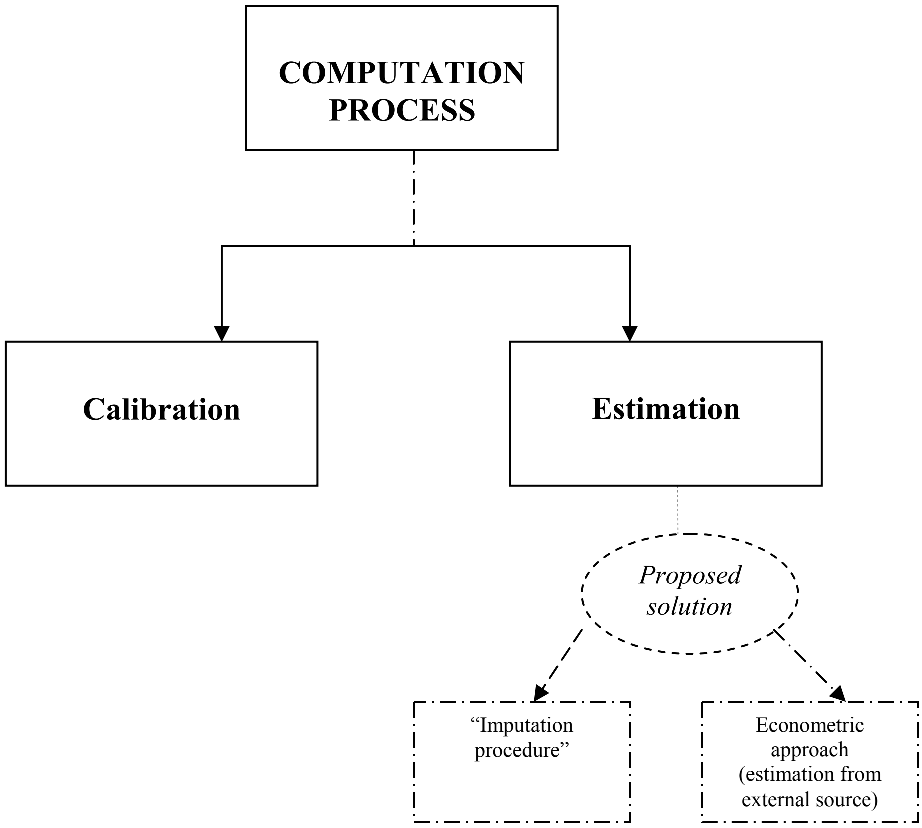
1.2 The Computation Process: Calibration versus Estimation?

The computation process of a RCGE model consists of both the calibration and parameter estimation. (Figure 1.1)
According to Dawkins et al. (2001) we will refer to calibration "as the setting of the numerical values of model parameters relative to a criterion of an ability to replicate a base case dataset as a model solution". To be precise the calibration refers to the use of information contained in a SAM in order to express macroeconomic equalities, identities of the model in quantitative terms.
On the other hand, we will refer to estimation as the use of a goodness of fit criterion in the selection of numerical values of model parameters.

In this respect, Devarajan and Robinson (2002) specified that the computation process of a CGE model basically requires two kinds of parameters:

i) share parameters, such as intermediate inputs costs, consumer expenditure shares, average saving rates, which can be determined from the RSAM which is the "core" dataset of a RCGE model under the well-known assumption that the base year represented by the RSAM is an equilibrium solution of the RCGE model. In this way, the RCGE follows the RSAM disaggregation of factors, activities, commodities and institutions;
ii) behavioural parameters and specifically elasticity parameters describing the curvature of various structural functions (such as production functions, consumption functions, import demand functions, export supply functions) which cannot usually be obtained from a single RSAM, but additional external data (secondary source of data) and external information are required for the estimation.

Figure 1.1. Computation process of a RCGE model: phases of calibration and estimation



The parameters at the point i) are determined in the calibration phase and therefore by using the information contained in the RSAM.
On the other hand, for recovering the parameter values specified at the point ii) it is not sufficient the information contained in the RSAM and therefore they are involved in the estimation phase.
For the "assignment" of these values two main procedures have been introduced in literature, the imputation and the econometric procedures, as better specified later.
The major debate on both RCGE and CGE models concerns this particular aspect of the computation process since the selection of appropriate parameters greatly influences, and in some cases even drives, the results of applied economic modelling exercises (Arndt et al. 2002; McDaniel and Balistreri 2002).

However, it is worth noting that there is a heated debate in literature aimed to define a borderline, if existent, between the calibration and the estimation phases of a RCGE model.
In this context, Hoover (1995) was one of the first authors to set out some of the lines of this debate discussing the empirical value of calibrated models over estimation. One of the most important paper on this issue is represented by the

research of Dawkins et al. (2001) in which they specified their position by the statement that "….calibration and estimation are in less conflict than one might suppose".
Balistreri and Hillberry (2005) - whose paper's purpose is to demonstrate that the statement of Dawkins et al. it is equally relevant to micro-based general equilibrium models – stressed that there is too little communication between calibrators and estimators of such models.
It is our intention in this research to enhance the already existing debate by demonstrating that these two phases in the RCGE framework, are really strictly related. In particular, by considering the computation process as a unitary process, in such a way that the two phases could be joined together, it is possible to parameterize all the unknown coefficients of production and demand functions characterizing a RCGE model, by only using the information contained in the RSAM.

1.3 Different Approaches to the Determination of the Behavioural Parameters

1.3.1 The imputation procedure

Generally, calibration consists in determining the numerical values of the various parameters of functions compatible with the equilibrium represented by the RSAM, within the RSAM itself.
However, in some cases information contained in a RSAM is insufficient or inadequate for the calibration of all parameters.
To be precise, when functional forms such as the CES, the Working-Leser, the Almost Ideal Demand System (AIDS) or the Linear Expenditure System (LES) are defined, the complete parameterization of these functional forms requires values of some additional parameters, namely the parameters describing the economic agent behaviour, such as the elasticity of substitution or the expenditure (or income) elasticity.
As mentioned before, the values of these parameters can be imputed or based on econometric estimations.

Referring to the first approach, known as the "*imputation approach*", the problem related to the behavioural parameter values is generally solved by imputing values taken from similar contexts and "borrowed" from existing studies. Moreover, where such estimations are not available for the country in question, assumptions on elasticity estimated for a country with similar characteristics can be applied. For example, this may mean using measures of central tendency of estimates (a kind of average) after carrying out a survey on existing literature.
This is a non-statistical approach as no estimation is performed and the approach is based on a criterion of analogy. It was greatly criticized by Jorgenson (1984), Lau (1984), Jorgensen et al. (1992), Diewert and Lawrence (1994), Nganou (2004) and other authors, since the values borrowed from literature can be more

appropriate for other countries rather than the country or the region the CGE model was intended for.
On this issue, Abdelkhalek and Dufour (1998) stress that the elasticity, available from literature are often distantly related to the case studied because they come from different countries or periods of times that those we are interested in.
The main consequence of working with this procedure is the great uncertainty regarding these basic ingredients, which is transmitted to the overall simulation results.
This procedure was also called "*calibration procedure*" by Mansur and Whalley (1984). These authors referred to the term "calibration" as a method for the behavioural parameter value selection and at the same time for indicating the "general" method which gives a numerical value to any parameter of a model. According to these Authors, the term calibration may concern both those parameters usually determined from the SAM (such as share parameters) and those parameters whose values cannot be obtained from the SAM and therefore are "borrowed" from existing studies.

In order to avoid confusion, from now on in this research we will refer to calibration only for the parameter values which can be determined from the SAM, while we will refer to the "borrowed" values from literature as the imputation procedure.

Despite the above-mentioned difficulties, many authors have taken advantage of this imputation procedure.
Piggott and Whalley (1980) used the substitution elasticity provided by Caddy (1976) - whose paper represents one of the most widely used sources in CGE computation - and the aggregate import and export CES price elasticity by Stern, Frencis and Shumacher (1976). These elasticities are used as point estimates to approximately compute the model at the benchmark equilibrium.
This approach was also adopted by De Melo and Robinson (1981), who obtained trade elasticity from Hickman and Lau (1973) and Alaouze (1977).
Shoven and Whalley (1992) did not estimate income and price elasticity, and used the imputation procedure as well (Ferrari and Manca, 2008)
Roland, Reinert and Shiells (1994) built a CGE model for three countries (USA, Canada and Mexico) and 26 sectors in order to analyze the impact of integration in North America and based their estimates on a study by Sobarzo (1992) for the same countries[4].
Defenders of the imputation approach pointed out that there are usually insufficient time-series data to reliably econometrically estimate the models. Shoven and Whalley (1984) tried to explain why the calibration approach was so widely used. In particular they underlined that some applied CGE models involved many thousands of parameters and the estimation of all of the model

[4] For the Armington elasticity for Mexico between domestic goods, imports from USA and Canada, and imports from the rest of the world the authors base their estimates on a study by Sobarzo (1992) for the same countries. These elasticities vary from 0.46 for "Other manufacturing" to 2.25 for "Agriculture".

parameters would require either unrealistically large numbers of observations or overly severe identifying restrictions using time series based estimation methods. At the same time benchmark data sets - the SAMs - are in value terms and the decomposition into separate price and quantity observations makes it difficult to sequence equilibrium observations with consistent units through time as would be required for time series estimation.

1.3.2 The econometric approach

As mentioned before, in the imputation approach it is common practice to "borrow" behavioural parameter values, such as elasticity, from other published data and on the basis of other researches. Although this procedure might sound straightforward, it is often exceedingly difficult because each study is different from every other and it is strictly necessary recognizing and taking account of these differences in the computation process of a CGE model (Shoven and Whalley, 1984). Moreover, as stressed by Adkins, Rickman and Hameed (2003) elasticity may have been estimated for geographical and industry classifications that are inconsistent with those required for the model in question.
Partridge and Rickman (2008) underlined that RCGE modellers routinely use the same external values of the behavioural parameters found in national or international models. In this way, these imputed values make the RCGE model inconsistent with the empirical evidence of the regional context.
Despite their extensive use, the "imputation" method has been quite criticized by, among others, Jorgenson (1984), Lau(1984), Jorgensen et al (1992), Diewert and Lawrence (1994) and McKitrick (1998) and on several grounds.
The main three features of these critiques are summarized in the work of McKitrick (1998) called by himself the "econometric critique" of CGE modeling.
First of all CGE modellers often make resort to elasticity (or in general to the behavioural parameters) estimated for commodity or industry classification which are inconsistent with those represented in the model and/or for countries other than the ones the CGE model studied is trying to represent. These expediencies detract from the ability of the model to represent "the technology and tastes of the economy under study" (McKitrick, 1998).
Secondly, a direct consequence of the imputation procedures is a partially dependence of the quality of the model constructed to the quality of the data used for the parameter values imputed.
Third, the imputation approach tends to limit the researchers to the use of "restrictive" functional forms all of which embody restrictive assumptions about the structure of the industries analyzed.
The first attempt to solve the difficulties related to the imputation procedure was provided by Jorgenson (1984) who presented an econometric model of producer behaviour suitable for incorporation into a CGE model. He implemented the econometric models for the producer behaviour and technical change by assembling a time series database (for the period 1958-1974) for thirty-six industrial sectors of the United States economy (Jorgenson and Fraumeni, 1981). His model was based on a production function for each sector, giving output as a

function of inputs of intermediate goods produced by other sectors and inputs of the primary factors of production, capital and labour services. Output also depends on time as an index of the level of technology. Producer equilibrium under constant returns to scale implies the existence of a sectoral price function, giving the price of output as a function of the input prices and time. He also incorporated the restrictions implied by economic theory of producer behaviour by generating a price function for each sector.
The most important conceptual innovation of Jorgenson's work is represented by the determination of the rate of technical change and the distributive shares of productive inputs simultaneously as functions of relative prices. Estimates of the unknown parameters of Jorgenson's econometric model of producer behaviour are based on the nonlinear three-stage least squares estimator introduced by Jorgenson and Laffont (1974).
A decade later, Mc Kitrick (1998) in his work has also specified that an empirical economic model, such as a CGE model, embodies three types of information: analytical, functional and numerical.
The analytical structure is the theoretical background material that identifies the variables of interest and their causal relations and for CGE models it is represented, usually, by the neoclassical canon.
The functional structure is the mathematical and quantitative representation of the analytical background and consists of the algebraic equations which make up the actual model.
The numerical structure consists of both the signs and magnitudes of the coefficients in the equations forming the functional structure.
McKitrick's econometric critique of CGE modelling is not directed to the analytical structure but it calls into serious questions the functional and numerical structures of the calibrated CGE models. To show that, he constructed two short-run CGE models in his paper; the first one is based on CES functional forms while the second model is based on normalized quadratic functions. In both models, all equations are econometrically estimated on a single 29-year time series database ad hoc constructed and whose data have been taken from Statistics Canada's Canadian Socio-Economic Information Management System.

To summarize the *econometric approach,* firstly introduced by Jorgenson (1984) and further developed by McKitrick (1998) refers to the use of time-series or cross-section data in addition to the SAM, in order to estimate the unknown behavioural parameters of a CGE model.
In this context an interesting application is provided by Arndt et al. (2002), who used time series data joined to prior information on the elasticity to perform an estimation of trade elasticity associated to a CES aggregated function, by using the entropy based estimation approach. In the model to be estimated, the data are aggregated up to six commodities (food, crash crops, processed food, fish, manufactures, services) and seven activities which correspond one to one to the

commodities plus the commerce activity[5]. Time series data of imports, exports, tariff revenue, total production, marketing margins, intermediate consumption and household consumption, referred to the period 1991-1996, are provided by the National Statistical Institute of Mozambico.
Adkins et al. (2003) used a Bayesian estimation approach to obtain production parameters for a RCGE model of Oklahoma manufacturing sector, in which regional data was sparse and judged to be of poor quality. The Bayesian priors formulated included information derived from the neoclassical restrictions and information from national data.
Liu, Arndt and Hertel (2003) presented a general approach to parameter estimation, based on an approximate likelihood function[6], and developed some goodness-of-fit measures for RCGE and global CGE models. The method is applied to estimation of Armington substitution elasticity in a relatively standard global model including 10 countries and focused on East Asian trade, employing a modified version of a standard, global CGE model developed by Rutherford (1999) and nick-named by themselves "GTAP in GAMS."
Nganou (2004) estimated own-price and income elasticities, as well as Frisch parameters for households, whose consumption behaviour is described by a LES demand function in the context of a CGE model for Lesotho. GME technique is used instead to estimate Armington parameters. Data used to estimate LES parameters consist of both 1986/87 Lesotho Household Expenditures Survey and 2000 Consumer Price Index series by commodities and location (rural and urban) and are provided by the Lesotho Bureau for Statistics.

However, the behavioural parameter econometric estimates from "*secondary source*", such as time series or cross section data in addition to the national or regional SAM, have been quite criticized. On one hand, by using external data in addition to the original dataset the introduction and the complete specification of more flexible functional forms would be possible but, on the other hand, this would create problems concerning level of aggregation and comparability of the data.
As argued by Blackorby and Russel (1989) the estimation of specific input parameters, such as elasticity, is affected by conditions and assumptions specific to each estimation process. More precisely the "external data" on which the estimates are based, creates several problems related to the comparability and reliability of the results since the aggregation level of the "secondary source" data might not be the same of the SAM under which the studied CGE model is constructed. In this way, it is worth noting the statement of Boys and Florax (2007) according to them the parameter estimates require thoughtful consideration of both the source of the data for estimation and the purpose for estimation and their utilization.

[5] Information on the SAM underlying the CGE model is available in Arndt, Cruz, Jensen, Robinson and Tarp (1998).
[6] They developed an approximate likelihood approach that focuses on discrepancies between model predictions and available data through time, across industries and across regions.

1.4 The Proposal of a “Self-Contained” Approach: Motivation and Main Contribution

The main objective of this thesis is to propose an estimation procedure which enables us to obtain the values of behavioural parameters by using the information contained in the RSAM only.
Our proposal moves from the existing approaches and in particular, from the faults and defects they have demonstrated to have.
By considering the economic framework in question, this approach will help to eliminate the problems connected both to the imputation procedure and to the econometric procedure originating from external sources.
As already mentioned several authors have often underlined how values “borrowed” from already existing studies do not necessarily adapt to a specific context of study and therefore they can bias the results of the RCGE model. In addition, a partial dependence of the quality of the model constructed to the quality of the data used for the parameter values imputed is closely connected to the imputation procedure.
The econometric procedure from external sources has also been widely criticized partly due to the unavailability of external data on which the behavioural parameters are estimated. In particular, the lower the level of analysis, the more difficult it will be to find the necessary data. On the other hand, even when this data is available it necessary to pay particular attention to the aggregation level since it can cause problems of comparability among the different sources of data and consequently problems of interpretation of the results.

As an alternative to the two above described main approaches, it is worth noting the approach introduced by Boys and Florax (2007) who have proposed to survey the existing literature and to combine published elasticity estimates by using econometric estimation method. Firstly, they constructed a database of elasticity estimates through an extensive literature review[7]. Then, they used a meta-regression analysis in order to identify structural sources of variation in elasticity estimates sampled from primary studies. In addition, in the research they underlined how the meta-regression analysis could be used to improve the estimation of these crucial economic parameters by combining relevant estimates, investigating the sensitivity of estimates to variations in underlying assumptions, identifying and filtering out publication bias, and explaining variation in reported estimates.

To 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 RSAM in order to estimate the behavioural parameters of production or demand functions.

[7] In order to identify the ensemble of studies available, they performed a comprehensive review of the agricultural production literature and in particular they referred to the agricultural input substitution elasticity between labour and capital.

The proposal of a "self-contained" approach for the estimation of key parameters, both for production and demand functions, which is an approach based only on the information contained in a RSAM, enables us to solve the above-mentioned problems, as it does not introduce any source of external data, in addition to the RSAM.

In this way, our approach is fully eligible to be included in the existing debate in literature aimed to define a "border-line", if existent, between the calibration and the estimation of a CGE model.

More precisely, our idea is to join the calibration and the estimation phases together, so that the computation process is only based on the information contained in the RSAM.

In actual fact, our approach is in total agreement with Dawkins et al. (2001) according to whom "calibration is estimation, estimation is calibration", and Balistreri and Hillberry (2005) for whom calibrators and estimators in reality communicate much less than they should and therefore this lack of communication leads to a lack of research.

Statistically speaking our approach leads to an "ill-posed" situation since a RSAM does not contain enough information to obtain valid estimates. In order to overcome this difficulty our approach is based on GME and GCE methods which refer to Shannon's entropy, Kullback-Leibler measure and the ME principle developed by Jaynes (1957a; 1957b). The methods used will be discussed in chapter 2.

Moreover, whenever we need to estimate production and demand function parameters it is necessary to specify the functional forms describing the economic agent behaviours. The choice of the functional forms for the Sardinian economic context will be carried out in Chapter 3 by considering both the theoretical aspects of the selected functional forms and the available data.

Chapter 2

Entropy Based Estimation Approaches

"...in making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known. This is the only unbiased assignment we can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have"
[Jaynes, 1957b]

2.1 Ill-posed Problems and Proposed Solutions

A large number of data sets, in economics and other social sciences, are non-experimental, very small, limited or highly collinear, that is to say there is not enough information for perfectly inverting the data matrix in order to obtain a solution (Golan et al. 1996a; Golan, 2003).
Mittelhammer, Judge and Miller (2000) stressed that the data on which the analyses are based, is often partial or incomplete, since one, or more, of the following situations can arise:

- repetitions of the experimental data generating process are unnatural if not impossible;
- the number of observations is limited and the unknowns may be greater than the data points;
- the covariates may be highly correlated;
- data is in the form of count data and it may be in the form of averages, sample moments, or other aggregate measures, where probabilities must be used to represent partial information about individual outcomes.

The same Authors suggest the use of the term "*ill-posed*" to refer generically to a situation characterized by any of the above-mentioned difficulties.

Fraser (2000) stresses that the following distinction has been made in literature in order to sum up the types of problems that can affect a dataset. A situation is described as "ill-posed" because of non-stationarity or because the number of observations is limited. Alternatively, a situation is defined as "ill-conditioned" or "ill-behaved" when the parameter estimates are highly unstable due to the bad design of the experiment or because the data is generated non-experimentally. One of the related problems in this case is the collinearity of data.

It is worth noting that in this research the term "ill-posed" will be used according to the general definition given by Mittelhammer et al. (2000), that can be referred to as "under-determined" as well.

Convenient assumptions, such as addition of prior or out of sample information to the original dataset, have been used to convert an "ill-posed" problem into a seemingly "well-posed" statistical problem. As stressed by Golan, Judge and Karp (1996c), ill-posed problems are often simplified by imposing "a priori" restrictions or by reducing the number of equations or unknowns, in order to use a traditional estimation method. Unfortunately, these devices often lead to unfavourable solutions and erroneous interpretations and treatments.
One of the current estimation methods specifically developed for dealing with these situations are the GME and GCE estimators, proposed by Golan et al. (1996a).
These estimation approaches lie on an extension of the entropy measure firstly introduced by Shannon (1948) and represent a generalization of the ME principle developed by E.T. Jaynes (1957a, 1957b).
As stressed by Arndt et al. (2002), the GME and GCE approaches have a great number of advantages for estimating parameters within the CGE or RCGE context. First, they allow to impose all general equilibrium constraints. Secondly, they permit incorporation of prior information on parameter values. Moreover, they can be applied in the absence of copious data.
The estimation of behavioural parameters of production and demand functions, in a RSAM-based approach is typically an "ill-posed" situation as in our case, since a RSAM does not contain enough information to obtain significant statistical estimates. Due to these peculiarities and taking into consideration the RCGE framework, the estimates will be carried out by taking advantage of these entropy-based approaches.

2.2 Entropy Criterion as Information Measure: an Historical Perspective

The origin of the term "entropy" dates back to XIX century and to thermodynamics, as a measure of disorder of a system.
More specifically, the second law of thermodynamics states that the entropy of a (closed) system (like the universe) increases with time. Thus, in this context, it represents the progression of a system towards equilibrium, that is reached at the highest level of entropy. Similarly, in the information theory field, entropy is defined as a measure of uncertainty or missing information.
In 1948, Claude E. Shannon, a communication engineer, was concerned with the problem of communicating information across noisy channels in which there was the potential for information loss during the communication process. One of his principal aim was to measure the level of uncertainty contained in a possibly noisy message received by an individual[1].

[1] Since the principal reason for providing an informational message is to remove or reduce uncertainty regarding some issue or topic, and since probability theory can be used to characterize uncertainty in a proposition or in the occurrence of events, Shannon proceeded to define a measure of uncertainty in probabilistic terms.

With this objective in mind and by using an axiomatic approach which delineated certain primitive and logical properties that a quantitative measure of information should possess[2], Shannon deduced a unique function, $H(p_1, p_2, ..., p_k)$, able to measure the degree of uncertainty contained in a message.
As the same author stressed in its seminal work, the entropy measure will quantify "…how much information is produced by such a process, or better, at what rate information is produced" (Shannon, 1948).
The introduced function, satisfying some specified assumptions, is of the form:

$$H(\mathbf{p}) = -\sum_{k=1}^{K} p_k \ln p_k \qquad [2.1]$$

(with the condition that 0 ln0=0) where the set of probabilities $\{p_1, p_{2,} ..., p_k\}$ are the probabilities linked to a discrete random value X with possible outcome values $\{x_1, x_{2,} ..., x_k\}$.

Each outcome has a probability $p_k=P(x_k)$, such that $\sum_{k=1}^{K} p_k = 1$.

In this way the measure's criterion defined by Shannon, H(**p**), called *entropy*[3], is function of the probability distribution **p** and not a function of the actual values taken by the random variable X.
Renyi (1961) argued that the entropy of a probability distribution can be interpreted not only as a measure of uncertainty but also as a measure of information. In actual fact, the amount of information, which we get when we observe the result of an experiment (depending on chance), can be taken numerically equal to the amount of uncertainty concerning the outcome of the experiment before carrying it out.
Intuitively, as stated by Mittelhammer, Judge and Miller (2000) information should be a decreasing function of p_i: the more unlikely an event, the more interesting it is to know that it can happen.
Golan (2008), in his survey and in the context of information theory[4], refers to entropy as expected information as well. The same Author states that entropy

---

[2] Shannon (1948) underlines that it is reasonable to require of it the following properties:
- i) H should be continuous in the p_i;
- ii) If all p_i are equal, $p_i=1/K$, then H should be a monotonic increasing function of n. With equally likely events there is more choice, or uncertainty, when there are more possible events;
- iii) If a choice be broken down into two successive choices, the original H should be the weighted sum of the individual values of H.

[3] As stressed by Golan (2008), Shannon in completing his work noted that "information" is already an overused term. The "legend" is that he approached his colleague John von Neumann, who responded: "You should call it entropy for two reasons: first, the function is already in use in thermodynamics under the same name; second, and more importantly, most people do not know what entropy really is, and if you use the word entropy in an argument you will win every time".

[4] The review and synthesis by Golan (2008) on information and entropy econometrics, discusses the concept of information and as it relates to econometric and statistical analyses of data. The meaning of "information" is studied and related to the basics of Information Theory as it is viewed by economist and researchers who are engaged in deciphering information from the data, while

reflects what we expect to learn from observations, on average, and it depends on how we measure information.

As stressed by Fraser (2000) the Shannon's entropy formulation coincides with the minus expectation of the logarithms of the probabilities and thus, it is considered as a measure of uncertainty or missing information.

In a more technical point of view, entropy is a measure of uncertainty of a single random variable and therefore it can be also viewed as a measure of uniformity. In such a way, Shannon's entropy measure can be also viewed as a measure of the distance between the discrete uniform distribution and the distribution generating **p**.

The values of Shannon's entropy range between *0* and *ln(k)* quantifying the degree of uncertainty in a distribution, **p**, measured on a scale ranging from a maximum to a minimum information level.

Entropy is minimized at the value of 0 if and only if **p** is degenerate on a particular outcome, x_j, so that $p_j=1$ and $p_i=0$ for $k\neq j$. The minimum value of entropy is associated with a maximally informative probability distribution for anticipating subsequent outcomes of the random variable, X, which implies that x_j will occur with probability equal to one. Thus, by using the statement of Shannon (1948), "…only when we are certain of the outcome does H vanish, otherwise H is positive".

In more detail, the minimal value of the Shannon entropy will be reached in a situation like the following:

$$H(\mathbf{p}) = -\sum_{k=1}^{K} p_k \ln p_k = -\left[0\ln 0 + 0\ln 0 + \ldots + 1\ln 1\right] = 0 \qquad [2.2]$$

On the other hand, the maximum value of entropy is uniquely associated with the maximally uninformative (or that is the same the minimal informative) weight-probability distribution. For a given K, H reaches its maximum when all *K* outcomes have the same probability to occur (i.e. equal to *1/K)*, that will be:

$$\begin{aligned} H(\mathbf{p}) &= -\sum_{k=1}^{K} p_k \ln p_k = -\left[\sum_{k=1}^{K} \frac{1}{K}\ln\frac{1}{K}\right] = -\left[\frac{1}{K}\sum_{k=1}^{K}\ln\frac{1}{K}\right] = \\ &= -\left[\frac{1}{K}\cdot K\ln\frac{1}{K}\right] = -\left[\ln 1 - \ln K\right] = \ln K \end{aligned} \qquad [2.3]$$

In a situation like that described by the [2.3] it can be noted that the statistical probability distribution associated to the maximum value of the Shannon entropy is the uniform distribution which treats all outcomes as equally likely and thus discriminate none of the potential outcomes as being more or less likely to occur than another.

---

taking into account what they know about the underlying process that generated these data and their beliefs about the economic system under investigation.

After the study of Shannon, other entropy measures have been developed. The most known are those introduced by Rényi (1961, 1970) and Tsallis (1988), which include the measure of Shannon as a special case.
Starting with the idea of describing the gain of information, Rényi (1961) has initially developed the entropy or order α for incomplete random variables (i.e. such that $\sum_{k=1}^{K} p_k \leq 1$) and later the generalized entropy measure of a proper probability distribution (Rényi, 1970) defined as:

$$H_\alpha^R(\mathbf{p}) = \frac{1}{1-\alpha} \ln \sum_k p_k^\alpha \qquad [2.4]$$

where the parameter α is assumed to be strictly positive (α>0).
The other information measure introduced by Tsallis (1988) can be defined as:

$$H_\alpha^T(\mathbf{p}) = c \frac{\sum_k p_k^\alpha - 1}{1-\alpha} \qquad [2.5]$$

where the value of c, a positive constant, depends on the particular units used.
Both of these measure[5] include the Shannon measure as a special case. Particularly as $\alpha \to 1$, $H_\alpha^R(\mathbf{p}) = H_\alpha^T(\mathbf{p}) = H(\mathbf{p})$.
It is interesting to note that, as stressed by Golan (2008), the functional form of Renyi's measure described by the [2.4], resembles the CES production function, while Tsallis's equation described by the [2.5] is similar to the Box-Cox function.

The three measures, introduced by the [2.1], [2.4] and [2.5] share the following three properties, as underlined by Golan and Perloff (2002):

i) they are nonnegative for any arbitrary **p**. In fact, the three entropy measures are always strictly positive except when all probabilities but one equal zero (which coincides to a perfect certainty information context);
ii) they reach a maximum value when all probabilities, related to all the possible events, are equal;
iii) each measure is concave for arbitrary **p**;

However, it is worth noting that in the Shannon's entropy events with high or low probability do not add much to the entropy level. Indeed, in the measures

[5] The Renyi and Tsallis entropies have been compared in Tsallis (1988) and Holste, Grobe and Herzel (1998) to show that:
$H_\alpha^R(\mathbf{p}) = (1/1-\alpha)\ln[1+(1-\alpha)\ln H_\alpha^T]$

introduced by the [2.4] and [2.5] higher probability events contribute more to the value than do lower probability events[6].
This property is one of the most important properties of the Shannon's entropy and represents, in the meanwhile, one of the strength points to the extension of this measure in a statistical inferential context.

Moreover, the three entropy measures differ in terms of their additive properties. In this context an important property, which distinguishes the Shannon's entropy from the other generalized measures proposed, is known as "*Composite Events property*".
This property states that the value of the Shannon's entropy for a composite event is equals to the sum of the marginal and conditional entropies.
To demonstrate this property it is firstly necessary to define two discrete and finite random variables X and Y, whose possible realizations are $\{x_1, x_2, ..., x_k\}$ and $\{y_1, y_2, .., y_j\}$, respectively. Let us define also the probability associated to each single outcome x_k and y_j as $P(X = x_k) = p_k$ and $P(Y = y_j) = q_j$; the joint probability $P(X = x_k, Y = y_j) = w_{kj}$; the conditional probabilities $P(X \mid Y) = P(X = x_k \mid Y = y_j) = p_{k|j}$ and $P(Y \mid X) = P(Y = y_j \mid X = x_k) = q_{j|k}$ where $p_k = \sum_{j=1}^{J} w_{kj}$, $q_j = \sum_{k=1}^{K} w_{kj}$ and $w_{kj} = q_j p_{k|j} = p_k q_{j|k}$.
The conditional entropy $H(\mathbf{p}|\mathbf{q})$ represents the total information contained in X with the condition that Y has a certain value (Golan, 2003):

$$H(\mathbf{p} \mid \mathbf{q}) = \sum_j q_j \left[-\sum_k p_{k|j} \log p_{k|j} \right] = \sum_j q_j \left[-\sum_k \left(\frac{w_{kj}}{q_j} \right) \log \left(\frac{w_{kj}}{q_j} \right) \right] = \\ = \sum_{k,j} w_{kj} \log \left(\frac{q_j}{w_{kj}} \right) \quad [2.6]$$

As shown by Renyi (1970) Shannon's entropy is the only measure which satisfies the following expression:

$$H(\mathbf{p}, \mathbf{q}) = H(\mathbf{q}) + H(\mathbf{p} \mid \mathbf{q}) = H(\mathbf{p}) + H(\mathbf{q} \mid \mathbf{p}) \quad [2.7]$$

and, if X and Y are two independent random variables the above equation reduces to:

[6]Golan (2003) stresses that in the measure [2.4] and [2.5],unlike the Shannon's measure, the average logarithm is replaced by an average of powers α. Thus, a change in α changes the relative contribution of the event k to the total sum. The larger the α, the more weight the "larger" probabilities receive the sum.

$$H(\mathbf{p},\mathbf{q}) = H(\mathbf{p}) + H(\mathbf{q}) \quad [2.8]$$

Another important property of the Shannon's entropy is the additivity, (also known as Shannon's additivity), according to which, the total amount of information in the entire K dimensional sample is equal to the weighted average of the information contained in two mutually exclusive sub-samples A and B of length L and M respectively (with L+M=K)[7].
The properties above-specified, together with further developments in econometric and statistical fields have conducted to view the Shannon's measure under different point of views, also in term of image reconstruction.
In the economic field, for example, at the beginning of the analysis a researcher never knows the specific true underlying values characterizing a system. Therefore, as underlined by Golan (2008), one may incorporate understanding and knowledge of the system in reconstructing (estimating) the image (represented by the unknown parameters) where this knowledge appears in terms of some global macro-level quantities such as moments. So, the entropy of the analyzed economic system measures the uncertainty of the researcher who knows only some moments' values representing the underlying population.
As stressed by Myung, Ramamoorti and Bailey (1996) Shannon's entropy together with mathematical communication theory has found significant applications in a variety of fields spanning astronomy, economics, engineering, geology, physics, statistics, transportation and urban and regional planning.

2.3 The ME Framework

2.3.1 The ME principle

In 1957, E.T. Jaynes formulated Shannon's entropy as a method for estimation and inference particularly for limited and insufficient data situations by proposing the so-called ME principle.
Jaynes (1957b) was motivated to develop his principle as a form of statistical inference, studying the prediction of equilibrium thermodynamic properties. His starting points were the Shannon's entropy as an information measure and the subjective interpretation of probabilities.
With these notions in mind, he developed the principle according to which "the probability distribution subject to whatever is known, provides the most unbiased representation of our knowledge of the state of the system". Moreover, the chosen

[7] Let the probabilities for the subsample A be $\{p_1,...,p_L\}$ and for the subsample B be $\{p_1,...,p_M\} = \{p_{L+1},...,p_K\}$ and define $p_A = \sum_{k=1}^{L} p_k$ and $p_B = \sum_{k=L+1}^{K} p_k$. Then, for all α (including α=1 which is the special case of Shannon's entropy),
$H(p_1,...,p_K) = H(p_A,p_B) + p_A H(p_1/p_A,...,p_L/p_A) + p_B H(p_{L+1}/p_B,...,p_K/p_B)$

probability distribution under the ME principle is the broadest one compatible with the available information.
As a general statistical inference method, the ME principle developed by Jaynes rests on the maximization of the Shannon's entropy subject to appropriate consistency relation (such as specifically moments of the studied data) and adding up constraints, as a basis for assigning or recovering the unknown probabilities **p** that characterize a given data set.
As underlined by Paris and Caputo (2001), following Jaynes, the idea of entropy defines a probability measure such that the probability distributions with high entropy are "favored" in the sense that they are more likely to occur and require less a priori information.
Obviously, the maximization of the entropy which is not subjected to any consistency constraint, but subject to the proper probability distribution constraint, yields to the maximum value for H(***p)*** and consequently the estimated distribution of the **p**'s returns to the uniform one which describes, as known, the state of complete uncertainty.
The introduction of some observed sample moments (such as means or variances) as constraints into the optimization problem takes the solution away from uniformity. Specifically, as underlined by Golan (2008), the more information there is in the data, the further away the resulting distribution is from uniformity or from a state of complete ignorance.
Mittelhammer et al. (2000) underline that the ME principle is based on an extension of LaPlace's "principle of insufficient reason" , on the basis of which one should choose the maximally uninformative uniform distribution as an estimate of the unknown probabilities vector, **p**, when there is maximal uncertainty about the values of the p_i's. In other words, the uniform distribution is the "right" distribution, when there is insufficient reason to choose any other distribution that favours some outcomes as being more likely than others[8].
The Jaynes' ME principle represents an extension of Laplace's Principle since it considers that moment information about **p** is available and it leads to choose the maximally uninformative uniform distribution as the available moment constraint information allows.
According to this approach, as stressed by Fraser (2000), two or more feasible distributions satisfying the constraints could be found. The one to be selected is the one which is least informative, or most uncertain according to the Shannon entropy criterion.
On this issue, Jaynes (1957a, 1957b), underlined that this procedure "…is the only assignment that one can make; to use any other would amount to arbitrary assumption of information which by hypothesis we do not have".

[8]As argued by Jaynes (1957a) the problem of specification of probabilities in cases where little or no information is available, is as old as the theory of probability. However, except in cases where there is an evident element of symmetry that clearly renders the events "equally possible", this assumption may appear just as arbitrary as any other that may be made. Since the time of Laplace, this way of formulating problems has been largely abandoned, owing to the lack of any constructive principle which would give us a reason for preferring one probability distribution over another in cases where both agree equally well with the available information. (Jaynes, 1957a) .

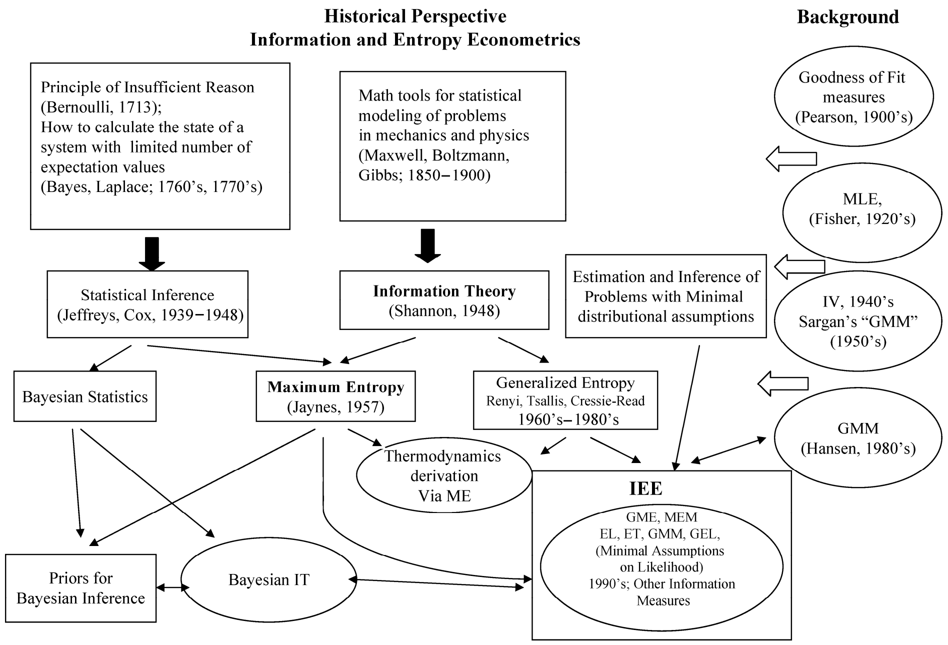
In such a way, it is possible to view the ME method as a method that yields the most uniformed distribution that is consistent with the observed sample moments. On this issue, Bera and Bilias (2002) underlines that maximization of Shannon's entropy subject to the data-moment constraints make the resulting estimated probabilities as smooth as possible.
The basic formulation of the ME principle proposed by Jaynes provides some new insight into the debate on "probabilities versus frequencies". In fact, he specified the notion of probability via Shannon's entropy measure. Jaynes' principle states that in any inference problem the probability should be assigned by this principle, which maximizes the entropy subject to the requirement of proper probabilities and any other available information, such as sample data which constitutes a constraint into the optimization problem.

2.3.2 ME principle and information entropy econometrics

An important aspect concerning the ME approach is the strictly correlation between the ME principle, view as an estimation method, and the information theory. Several authors in literature have investigated this relation.
Kesavan and Kapur (1989) underlined that the ME principle draws together concepts from information theory, statistical inference, optimization and "last but non least" a precise knowledge of the partial information one has about a probabilistic system in terms of a set of statistical moments.
Fraser (2000) underlines that in the information theory, as well as for the ME principle, the information contained in an observation is inversely proportional to its probability.
Golan (2008) has recognized two main and parallel paths that lead to the classical ME through the Information Theory (IT). These two lines of research, as the author underlines, are similar. The objective of the first line, pertaining to the 18th century and identified by the works of Bernoulli, Bayes and Laplace, is to formulate a methodology that allows understanding of the general characteristics of a system from partial and incomplete information. In the second line of research, pertaining to the 19th century and identified by the studies of Maxwell and Boltzmann and continued by Gibbs and Shannon (1948), this same objective is expressed as determining how to assign numerical values of probabilities when only some theoretical limited global quantities of the studied system are known.
In the light of these two research paths, it is possible to view the ME formalism as a principle based on the philosophy of the first line of research and the mathematics of the second line of research.
In a more general fashion, the ME principle represents one of the most important starting point of the Information and Entropy Econometrics (IEE) which is, as underlined by Golan (2007), the sub-discipline of processing information from limited and noisy data with minimal a priori information on the data-generating process. (Figure 2.1).

The main relationship linking all estimation methods which are included in the IEE[9] is that, rather than starting with a pre-specified likelihood, they take advantage of the observed data, in order to estimate a set of natural weights, or empirical distribution, which is most consistent with the observed sample moments or data (Golan, 2007).

Figure. 2.1 IEE: historical perspective



Source: Golan, 2008

Estimates obtained by using these methods are based on a maximization or minimization problem of a certain information criterion subject to the observed sample information. The objective functions used in these optimization problems are related to entropy of Shannon, which is a special case of the generalized entropy measure. (Golan, 2008). [10]

---

[9] This class of estimators includes the Empirical Likelihood (EL), the Generalized EL (GEL), the GME and the Generalized Method of Moments (GMM), as well as the Bayesian Method of Moments (BMOM). Golan (2008) provides a detailed discussion about this class of estimators.

[10] The core objective of all research within the IEE framework can be found in a simple example contained in the work of Jaynes (1963) and concerning a die toss experiment. In this example it is known the empirical mean value of a very large number of tosses of a six-sided die. With this information we wish to predict the probability that in the next throw of the die we will observe the value 1,2,3,4,5 or 6. Under the assumption that the probability vector is proper, we should predict six values (*unknowns*) with only two observed moments (*knowns*): the mean and the sum of probabilities. Since there are more unknowns than knowns there are infinitely many probability distributions that sum up to one and satisfy the observed mean. But, as stressed among other by Golan (2007), "which one of the infinitely many solutions should one use?". In all of the IT methods, the one solution chosen is based on either the Shannon's information measure-entropy, or other information measures that are directly connected to Shannon's entropy.

The ME principle developed by Jaynes is specifically designed to deal with ill-posed situations and as an estimation method able to recover a probability distribution from an insufficient number of moments which represent the only available information.
In fact, by using this approach the “chosen” probability distribution for the data is the one, which maximizes entropy and which is at the same time the least informative over the set of probability distributions that are consistent with the evidence.

2.3.3 Structure and solution of a ME problem

As Jaynes (1984) noted, when ill-posed problems arise, creative assumptions or prior information have been used in literature to induce a well-posed problem which lead to feasible solutions. As an alternative to this, the same author introduced, as mentioned above, the ME principle specifically oriented to deal with these types of problems and based on the measure introduced by Shannon.
The use of this formulation makes possible the transformation of the observed evidence, represented by the data in form of such empirical moments, into a distribution of probability describing our state of knowledge.
In this way, it is possible to convert, as stressed by Golan et al. (1996a) the problem from one of deductive mathematics to one of inference involving an optimization procedure, where “we seek to make the best, most objective predictions possible for the information that we have” (Mittelhammer et al, 2000).
The philosophy underlying this approach is to use *all*, and *only*, the information available for the estimation problem at hand.
Assuming that the ill-posed inverse problem might be written as:

$$\mathbf{y} = \mathbf{Xp} \qquad [2.9]$$

where $\mathbf{y}$ is a T-dimensional vector of observed moments (i.e. average or aggregated data) and $\mathbf{X}$ is a (T×K) design matrix. Our objective is to estimate the K-dimensional (with K>T) proper probability distribution $\mathbf{p}$.
According to the standard literature, (Golan et al, 1996a; Golan, 2008) the analytical definition, structure and solution of the estimation problem can be illustrated as follows:

$$\max_{p} H(\mathbf{p}) = -\sum_{k=1}^{K} p_k \ln p_k \qquad [2.10]$$

subject to moment-consistency constraints:

$$y_t = \sum_k x_{tk} p_k \qquad \text{t=1,2,...T} \qquad [2.11]$$

and to the adding-up normalization constraints defined on the probabilities $\mathbf{p}$, such that:

$$\sum_k p_k = 1 \qquad [2.12]$$

To recover the unknown probability vector **p** one can form the Lagrangian function:

$$L = -\sum_{k=1}^{K} p_k \ln p_k + \sum_{t=1}^{T} \lambda_t \left[y_t - \sum_{k=1}^{K} p_k x_{tk} \right] + \mu \left(1 - \sum_k p_k \right) \qquad [2.13]$$

with the first-order conditions:

$$\begin{aligned}
&\frac{\partial L}{\partial p_k} = -\ln \hat{p}_k - 1 - \sum_{t=1}^{T} \hat{\lambda}_t x_{tk} - \hat{\mu} = 0 \qquad \text{k=1,2,..., K} \\
&\frac{\partial \mathrm{L}}{\partial \lambda_t} = y_t - \sum_{k=1}^{K} \hat{p}_k x_{tk} = 0 \qquad \text{t=1,2,...,T} \\
&\frac{\partial L}{\partial \mu} = 1 - \sum_{k=1}^{K} \hat{p}_k = 0
\end{aligned} \qquad [2.14]$$

The solution to this system of K+T+1 equations and parameters yields

$$\hat{p}_k = \frac{\exp\left(-\sum_t \hat{\lambda}_t x_{tk} \right)}{\sum_k \exp\left(-\sum_t \hat{\lambda}_t x_{tk} \right)} \equiv \frac{\exp\left[-\sum_{t=1}^{T} \hat{\lambda}_t x_{tk} \right]}{\Omega\left(\hat{\boldsymbol{\lambda}} \right)} \qquad [2.15]$$

where $\Omega\left(\hat{\boldsymbol{\lambda}}\right) = \sum_k \exp\left(-\sum_t \hat{\lambda}_t x_{tk} \right)$ is a normalization factor. In particular, the factor Ω converts the relative probabilities to absolute probabilities and it is known in literature as the partition function.

The ME approach, in the form proposed by Jaynes (1957a; 1957b) has been primarily applied in the solution of estimation problems when information is given in the form of moment constraints or other aggregate measures.

However, as argued by Jaynes (1982), the ME can be applied to probability estimation problems involving any other constraints, not necessarily in the form of moments, yielding distinct maximum entropy solutions.

Golan et al. (1996a) generalized the traditional formulation proposed by Jaynes (1957a; 1957b) by defining the ill-posed pure inverse problem as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} \qquad [2.16]$$

where $\mathbf{y}$ is a T-dimensional vector, $\mathbf{X}$ is a (T×K) design matrix composed of explanatory variables and $\boldsymbol{\beta}$ is a K (with K>T) dimensional vector of unknown parameters.
Mittelhammer et al. (2000) refer to the term *pure* to point out the absence of the error term (the noise component) in the relationship between variables and parameters; that is the relationships are assumed to be exact.
In the aim of specifying the optimization problem on which the ME formulation proposed by Golan et al (1996a) is based, it is firstly necessary to re-parameterize each unknown parameter β_k in terms of probabilities because of the fact that the arguments of the Shannon's entropy, as mentioned above, are probabilities.
Particularly, it is necessary to introduce a set of equally distanced discrete points $\mathbf{z} = (z_1, z_2, ..., z_M)'$ with corresponding probabilities $\mathbf{p_k} = (p_{k1}, p_{k2}, ..., p_{kM})'$ and with M≥2 where M is the number of support points defined on each β_k.
Consistent with this specification it is possible to rewrite $\boldsymbol{\beta}$ as:

$$\boldsymbol{\beta} = Z\mathbf{p} \tag{2.17}$$

where:

$$Z\mathbf{p} = \begin{bmatrix} \mathbf{z}' & & & \\ & \mathbf{z}' & & \\ & & \cdots & \\ & & & \mathbf{z}' \end{bmatrix} \begin{bmatrix} \mathbf{p_1} \\ \mathbf{p_2} \\ \cdots \\ \mathbf{p_k} \end{bmatrix} \tag{2.18}$$

Z is a block diagonal (KxKM) matrix of support points with:

$$\mathbf{z}'\mathbf{p_k} = \sum_{m=1}^{M} z_m p_{km} = \beta_k \quad \text{for k=1,2,...,K} \quad \text{m=1,2,...,M} \tag{2.19}$$

By doing so, the re-parameterized system is described as:

$$\mathbf{y} = \mathbf{XZp} \tag{2.20}$$

As stressed by Eruygur (2005) by taking advantage of this re-parameterization, each parameter is converted from the real line into a well-behaved set of proper probabilities defined over the supports. The implementation of a ME estimation problem makes possible to choose and define for the unknown parameter a set of discrete points, which is called support space, on the basis of knowledge about size and magnitude of the unknowns. As a consequence of this, the specified support spaces could be different for each of the unknown parameters.
However, in most cases researchers have not any knowledge about the unknown parameters and therefore, as stressed by Golan et al. (1996a) they should specify for each of the unknown β_k a support space that is uniformly symmetric around

zero with end points of large magnitude, say $\mathbf{z}'_k = (-C, -C/2, 0, C/2, C)$ for M=5 and for some scalar C.
The equation [2.20] as introduced above for the data-moment formulation, will enter the optimization problem as data-consistency constraints. Also in this case, the related optimization problem consists in the maximization of the Shannon entropy measure for the H(**p**) subject to the constraints represented by the data and the normalization contraints introduced on the unknown probability **p**. The solution of the maximization problem yields to the $\hat{\boldsymbol{\beta}}$ expressed as: $\hat{\boldsymbol{\beta}} = Z\hat{\mathbf{p}}$ and in scalar form as $\widehat{\beta}_k = \sum_m z_m \widehat{p}_{km}$.

2.3.4 The CE approach

Kullback and Leibler (1951) and Kullback (1959) have introduced a more general measure of entropy specifically concerning the measure of the distance between two probability distributions, **p** and **q**. In general, as underlined by Preckel (2001) cross entropy measures the discrepancy between two distributions.
This measure, called cross-entropy (CE) or the Kullback-Leibler information criterion, can be expressed as follows:

$$I(\mathbf{p},\mathbf{q}) = \sum_{k=1}^{K} p_k \ln(p_k / q_k) \qquad [2.21]$$

As stressed, among others, by Mittelhammer et al. (2000) the CE formulation is a one-to-one function of Shannon's entropy described by the [2.1] when **q** is a discrete uniform distribution. Under this point of view, this entropy formulation allows for prior information on the probability distribution values **p**, pertaining to the data, in the form of a proper probability distribution **q** which describes some form of prior conceptual knowledge.
This general entropy formulation can be translated in the inference estimation context, by formulating the optimization problem and illustrating the related solution.
In this way, as underlined by Fraser (2000) it is possible to derive an estimate of **p** subject to the existing constraints, in such a way that it can be discriminated from **q** with a minimum of difference.
A more detailed description about the optimization problem connected to the CE estimator can be presented as follows.
Firstly, it is necessary to suppose that, in addition to the data, some source of additional information concerning prior knowledge on the unknown parameters is available. This additional available information is expressed in the form of a proper probability vector **q**.
The objective of the optimization problem may be reformulated into minimize the entropy distance between the data in the form of **p** and the prior **q**.[11].

[11] That is, the underlying principle is that of probabilistic distance or divergence.

Good (1963) refers to this objective as the minimization of the cross-entropy between the probabilities that are consistent with the information in the data and the prior information **q**.
The solution of the optimization problem is reformulated into find for **p**, out of all the distribution satisfying the one closest to **q**.
Following the metric introduced for the classical ME, the CE criterion between **p**, and **q**,can be expressed in the following way (Fraser, 2000):

$$I(\mathbf{p},\mathbf{q}) = \sum_{k=1}^{K} p_k \ln\left(\frac{p_k}{q_k}\right) = \sum_{k=1}^{K} p_k \ln p_k - \sum_{k=1}^{K} p_k \ln q_k \qquad [2.22]$$

Starting from a pure inverse problem like the one described by the [2.16], the related optimization problem, - where $\mathbf{p_k} = \left(p_{k1}, p_{k2}, ..., p_{kM}\right)'$ and $\mathbf{q_k} = \left(q_{k1}, q_{k2}, ..., q_{kM}\right)'$ represent the vector of unknown (to be estimated) and prior (known) probabilities respectively - can be specified by:

$$\min_{p} I(\mathbf{p},\mathbf{q}) = \sum_{k}\sum_{m} p_{km} \ln\left(\frac{p_{km}}{q_{km}}\right) \qquad [2.23]$$

subject to the data-constraints:

$$\sum_{k}\sum_{m} x_{tk} z_m p_{km} = y_t \qquad \text{for } t=1,2,...,T \qquad m=1,2,...,M \qquad [2.24]$$

and the normalization adding up constraints.

$$\sum_{m} p_{km} = 1 \qquad \text{for all m} \qquad [2.25]$$

The solution of the optimization problem leads to the $\widehat{\boldsymbol{\beta}}$ which in scalar form can be expressed as:

$$\widehat{\beta}_k = \sum_{m} z_m \widehat{p}_{km} . \qquad [2.26]$$

It is important noting that with an uninformative prior, i.e. uniform **q**'s, the maximum entropy solution returns. In such a sense, the ME can be viewed as a special case of CE.

2.4 GME and GCE Estimation Approaches

2.4.1 From the ME to the GME estimator

Although the ideas of ME and CE have entered the economic and econometric literature, there have been difficulties in employing the concepts in their traditional form due to some technical difficulties. In particular, Fraser (2000) pointed out that the ME and CE approaches, in their primal version, by assuming the moment condition in the constraint sets, do not specifically account for the presence of disturbances. Golan (2008) stresses that in these situations the moment conditions have to be exactly fulfilled (zero-moment conditions) and that this property could be satisfactory for relatively large samples or for well-behaved samples

Concerning the economic system, the ME and CE approaches focus on relations that can be specified trough pure inverse problems. Unfortunately, in both the social and economic sciences we have to face up to small and/or ill-behaved data where the moment's restrictions may be too costly. Particularly in the economic context, processes are typically interdependent, dynamic and stochastic, and the available economic data is often composed of limited and/or non-experimental observations.

Given these characteristics of both the nature of the economic processes and data, few observations are without measurement errors and few economic relations are free of shocks.

In the presence of these difficulties and by introducing the disturbance terms, the task of information recovery becomes, as argued by Golan et al. (1996a) an *inverse problem with noise*.

These authors, with these considerations in mind, have introduced in the early 1990's the GME and the GCE estimators.

In particular, the GME estimator uses a more flexible set of moment conditions in the optimization.

As argued by Golan (2008), this method provides a greater flexibility resulting in more stable estimates for finite and/or ill-behaved data and provides the researcher with a general framework for incorporating economic theoretic and other behavioural information in a simple way that is consistent with information theory[12]. Generally speaking, this estimator, by treating the moments (and particularly each observation as we will specify shortly) as stochastic, can be viewed as the classical ME reformulated with stochastic moment conditions.

Golan (2008) defines the term "stochastic moments" as moment conditions, or functions of the random variables and the unknown parameters[13], with additive terms that have expectation of zero.

[12] This information can be in terms of linear, non linear or inequality functions and does not have to be formulated in terms of zero-moment functions.

[13] These moments, or functions, can be conditional or unconditional. The stochastic moments can be introduced in two ways. First, by allowing for some additive noise (with mean zero) for each one of the moments conditions. Second, by viewing each observation as a noisy moment resulting from the same data generating process.

By doing so, in the GME approach each observation can be treated as a composite of two components: the signal, represented by the data and the noise, represented by the disturbance component.
The optimization problem under which the GME estimator is based, will be discussed in the next pages.

2.4.2 Structure and solution of a GME problem

The optimization problem on which GME estimator is based, refers to the entropy measure introduced by Shannon (1948) and represents, in the meanwhile, a generalization of the ME principle developed by Jaynes (1957a, 1957b).
To explain how GME works, let us consider the standard linear regression model, represented in matrix form, with T observation and K explanatory variables:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \qquad [2.27]$$

where $\mathbf{y}$ is a T dimensional vector of noisy observation, $\mathbf{X}$ is a (T×K) design matrix composed of explanatory variables and $\boldsymbol{\beta}$ is a K-dimensional vector of the unknown parameters which we want to recover from the data; $\mathbf{e}$ is a T-dimensional vector of the unobserved and unobservable disturbances.
In the GME approach, rather than search for the point estimates $\boldsymbol{\beta}$, each β_k is viewed as the mean value of some well defined random variable z. To explain this important notion it is firstly necessary to express each β_k as a discrete random variable with a compact support and $2 \leq M \leq \infty$ potential outcomes, where M indicates the number of elements in the specific support.

Let z_k be an M-dimensional vector $\mathbf{z_k} \equiv \left(z_{k1},...,z_{kM}\right)'$ for all k=1,2,3,…,K., where z_{kl} and z_{kM} represent, the lower and upper bounds on the support of each β_k respectively. By defining $\mathbf{p}_k$ as an M-dimensional proper probability distribution defined on the set $\mathbf{z}_k$ it is possible to express the k-th parameter as a convex combination of points $\mathbf{z}_k$ with weights $\mathbf{p}_k$ it is possible to formulate each β_k as:

$$\beta_k = \sum_m p_{km} z_{km} \equiv E_{pk}\left[\mathbf{z_k}\right] \qquad [2.28]$$

Without loss of generality, following Golan et al. (1996a) for the simple case M=2, each β_k can be expressed as:

$$\beta_k = p_k z_{k1} + \left(1 - p_k\right) z_{kM} \qquad [2.29]$$

This procedure has to be done for each unknown coefficient and further by assembling in matrix form, any $\boldsymbol{\beta}$ could be written as:

$$\boldsymbol{\beta} = \mathbf{Zp} = \begin{bmatrix} \mathbf{z}_1' & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{z}_2' & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{z}_K' \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ . \\ \mathbf{p}_K \end{bmatrix} \qquad [2.30]$$

where **Z** is a K×(K×M) support discrete diagonal matrix of block diagonal matrix i=1,…,N and $\mathbf{p} \gg \mathbf{0}$ is a KM-dimensional vector of "unknown" and further estimated weights. The weights have to be strictly positive and to sum to 1 for each *k*.

In this formulation, the observed data, **y**, are viewed as the mean process **Z** with a probability distribution **P,** that is defined on the supports z_k's and is conditional on **X**. Thus, as pointed out by Golan (2008) the choice of the support space is useful to estimate the P's which yield the point estimates (β).

The unobserved error vector **e** is also viewed as another set of unknowns, and similar to the signal vector β, each e_t is treated as a finite and discrete random variable with *J* possible outcome ($2 \leq J < \infty$). Specifically, as already done with the β's above, each error term is redefined as:

$$e_t = \sum_j w_{tj} v_j \equiv E_{w_t}\left[\mathbf{v}\right] \qquad [2.31]$$

where the index *t* refers to the number of observations.

The vector **e** of unknown disturbances is re-parameterized as well and it may be written in matrix form as:

$$\mathbf{e} = \mathbf{Vw} = \begin{bmatrix} \mathbf{v}_1' & \mathbf{0} & . & \mathbf{0} \\ \mathbf{0} & \mathbf{v}_2' & . & \mathbf{0} \\ . & . & . & . \\ \mathbf{0} & \mathbf{0} & . & \mathbf{v}_T' \end{bmatrix} \begin{bmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ . \\ \mathbf{w}_T \end{bmatrix} \qquad [2.32]$$

where the matrix **V** is a (T×TJ) support discrete block diagonal matrix and **w** is a TJ-dimensional vector of weights. With this re-parameterization also the error terms can be viewed as elements taken as random draws from a certain distribution with probability weights w_{tj}.

For both the unknown parameters and the error terms, support spaces are constructed, here and later with reference to production and demand functions in a RCGE context, as discrete and bounded. However, it could be possible to specify, within the same framework, unbounded and continuous supports. Golan and Gzyl (2002; 2003; 2006) specified, in detail, this type of supports.

Similar to other IT estimators discussed earlier, the main objective of the GME method is to estimate the unknown β with minimal distributional assumptions.

The starting point of a GME optimization problem, as illustrated above, is the specification of a finite and discrete support for each-one of the unknown parameters and for the error component.
As stressed by Golan et al. (2001), the GME method carries out a re-parameterization and a re-formulation of a general linear model (GLM) in order to estimate the parameters inside the framework of the extended ME principle.
In this way, the GME approach uses a flexible, dual-loss objective function: a weighted average of the entropy of the deterministic part of the model and the entropy from the disturbance or stochastic part.
Thus, the objective is to estimate simultaneously the full distribution for each β_k and each e_t with minimal distributional assumptions.
Using the above specified re-parameterizations, Judge and Golan (1992), firstly, in their earlier work rewrite the standard linear regression model as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} = \mathbf{XZp} + \mathbf{Vw} \qquad [2.33]$$

where vectors **p** and **w** represent, as well, the dose of a prior information on both parameter’s support space and disturbances.

GME estimator is obtained by maximizing the joint entropies of the distributions of the coefficients and the error terms subject to the data and the requirements for proper probabilities. In scalar notation, the GME formulation for a noisy inverse problem may be expressed as:

$$\max_{\mathbf{p},\mathbf{w}} H(\mathbf{p},\mathbf{w}) = -\sum_{k=1}^{K}\sum_{m=1}^{M} p_{km} \ln p_{km} - \sum_{t=1}^{T}\sum_{j=1}^{J} w_{tj} \ln w_{tj} \qquad [2.34]$$

subject to the constraints:

$$y_t = \sum_{k=1}^{K}\sum_{m=1}^{M} x_{tk} z_{km} p_{km} + \sum_{j=1}^{J} v_{tj} w_{tj} \qquad \text{for t=1,2,...,T} \qquad [2.35]$$

$$\sum_{m=1}^{M} p_{km} = 1 \qquad \text{for k=1,2,...,K} \qquad [2.36]$$

$$\sum_{j} w_{tj} = 1 \qquad [2.37]$$

The solution to this maximization problem is unique. Forming the Lagrangean and solving for the first-order conditions yields the optimal solution for **p** and **w**.
Each estimated $\widehat{p}_{km}$ and $\widehat{w}_{tj}$ is expressed as:

$$\widehat{p}_{km} = \frac{\exp\left(-\sum_{t=1}^{T}\widehat{\lambda}_t z_{km} x_{tk}\right)}{\sum_{m=1}^{M}\exp\left(-\sum_{t=1}^{T}\widehat{\lambda}_t z_{km} x_{tk}\right)} \equiv \frac{\exp\left(-\sum_{t=1}^{T}\widehat{\lambda}_t z_{km} x_{tk}\right)}{\Omega_k\left(\widehat{\lambda}_t\right)} \qquad [2.38]$$

where $\Omega_k\left(\widehat{\lambda}_t\right) = \sum_{m=1}^{M}\exp\left(-\sum_{t=1}^{T}\widehat{\lambda}_t z_{km} x_{tk}\right)$;

$$\widehat{w}_{tj} = \frac{\exp\left(-\widehat{\lambda}_t v_{tj}\right)}{\sum_{j=1}^{J}\exp\left(-\widehat{\lambda}_t v_{tj}\right)} \equiv \frac{\exp\left(-\widehat{\lambda}_t v_{tj}\right)}{\Omega_t\left(\widehat{\lambda}_t\right)} \qquad [2.39]$$

where $\Omega_t\left(\widehat{\lambda}_t\right) = \sum_{j=1}^{J}\exp\left(-\widehat{\lambda}_t v_{tj}\right)$

The estimated probabilities $\widehat{p}_{km}$ and $\widehat{w}_{tj}$ produces the GME point estimates and the full distribution of β_k and e_t as:

$$\widehat{\beta}_k = \sum_{m=1}^{M}\widehat{p}_{km} z_{km} \qquad \text{for k=1,2,...,K} \qquad [2.40]$$

$$\hat{e}_t = \sum_{j=1}^{J}\widehat{w}_{tj} v_{tj} \qquad \text{for t=1,2,...,T} \qquad [2.41]$$

As stressed by Eruygur (2005) and other authors, the ***β***'s GME estimates depend on the optimal Lagrange multipliers $\widehat{\lambda}_t$ for the model constraints. These multipliers reflect the "marginal information" of each observation.

Golan (2008) outlines, as well, that like in the Empirical Likelihood (EL) the Lagrange multipliers capture the *natural weight* of each observation and transmit that information in the estimated exponential distribution $\widehat{p}_{km}$ and $\widehat{w}_{tj}$.

Particularly, within the GME the "natural weight" of each observation (defined also the "natural empirical distribution") is just a function of the T Lagrange multipliers:

$$\pi_t(\widehat{\lambda}) = \frac{\exp(-\widehat{\lambda}_t)}{\sum_t \exp\left(-\widehat{\lambda}_t\right)} \qquad [2.42]$$

As $T \to \infty, \pi \to$ uniform distribution. Unlike the other IT estimator, under the GME method, the weights are direct function of the information in the data. Specifically, the estimated Lagrange parameters reflect the contribution (in terms of information) to the optimal level of the entropy (objective) and as such, they capture the contribution of each observation to the "explanation" of the signal.

There is no-closed form solution for $\hat{\lambda}_t$ and hence no closed form solution for **p**, **w**, **β** and **e**. Therefore numerical optimization techniques should be used to obtain the solutions and solutions must be found numerically.

As stressed, among the others, by Golan (2003) the main difference between the "classical" ME approach and the GME is that in the former, data is in terms of moments (and specifically pure moments) while in the latter each (noisy) observation is taken directly into account.

Nganou (2004) outlines as well that ME is a special case of the GME where no weight is placed on the entropy of the error terms and where data is represented in terms of exact moments.

Thus, the GME method does not require a regularization parameter or a-priori assumptions on the exact nature of the relationship between the observed sample moments and the unobserved moments of the population. Instead, the regularization appears through the pre-specified bounds on the supports **Z** and **V** that re-parameterize both the unknown parameters and the error terms.

By these re-parameterizations the solution values of the parameters are consistent with the prior beliefs about the range of plausible parameter values and with the available sample.

2.4.3 The GCE estimator

Additional information about the unknowns parameters may be mainly and easily expressed in this framework by two ways:

i) as already discussed above, and as we will specify in a greater detail in the following pages, it is possible to incorporate information in the form of specifying the upper/lower bounds on a support space for each unknown parameter;

ii) in addition "a priori" information about the unknowns may be expressed in the form of prior probability distributions on the supports concerning both the unknown parameters and the error terms.

The first situation represents the already above-specified situation and it is included within the GME approach.

Instead, the other situation leads to a "transformation" of the GME approach into the GCE estimator. The introduction of prior-beliefs about the unknown parameters by using a probability distribution is the main difference between the GME and the GCE approach.

As already specified for the classical CE approach in the GCE, specifically, the unknown parameters **p** and **w** are subject to prior information, expressed as a probability distribution, **q** and **u**.

Particularly, the GCE formulation minimizes the entropy measure between prior assessments of a parameter and the estimated value. If in the GME formulation one maximizes the entropy, under the GCE formulation the entropy difference between **p** and **q** is minimized. If GCE yields a value greater than zero then the sample data have yielded a gain in information and learning can be assumed to have occurred. With repeated samples, GCE is a form of shrinkage rule so that the constructed probability approaches the true probability as the sample size approaches infinity (Fraser, 2000)

GME and GCE methods are similar in the sense that the GME can be viewed as the GCE with uniform **q**'s. In this way, the optimization GME problem can be also viewed as a minimization of the joint entropy distance between the data and the state of complete uncertainty, namely the uniform distribution.

By using the GCE method, the entropy objective is used to find the set of "posterior" distributions on the supports that satisfy the observations and are "closest" to the prior distributions[14]. Specifically, the GCE approach proceeds by minimizing the entropy between a prior estimate and the reconstructed probability.

Transforming GME into GCE formulation yields the following objective function, with ($\mathbf{p}, \mathbf{w} \gg 0$):

$$\begin{aligned} \min I(\mathbf{p},\mathbf{q},\mathbf{w},\mathbf{u}) &= \mathbf{p}'\ln(\mathbf{p}/\mathbf{q}) + \mathbf{w}'\ln(\mathbf{w}/\mathbf{u}) = \\ &= \sum_m \sum_k p_{km} \ln p_{km} - \sum_m \sum_k p_{km} \ln q_{km} + \sum_j \sum_t w_{tj} \ln w_{tj} - \sum_j \sum_t w_{tj} \ln u_{tj} \end{aligned} \qquad [2.43]$$

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

In general GME/GCE solutions behave like other shrinkage estimators, the variance of the estimator is less than the variance of the sample-based rules like Least Squares (LS) or Maximum Likelihood (ML), but the use of prior information introduces bias. This bias is typically offset by variance reductions – so the mean squared error of the estimator is smaller than sample-based mean squared error.

The estimation problems that will be carry out in this research, as further and better explained, will be based on the GME and the GCE approaches.

In particular, we will implement a three step entropy approach to estimate the production function parameters. Concerning the demand system, we will take advantage of the GME property, as well.

Estimates will be obtained by using the software GAMS (General Algebraic Modeling System) and the Minos and Path non linear solvers.

---

[14] As stressed by Golan et al. (1996a) there are significant similarities between the entropy solutions and the posterior resulting from Bayes' rule.

2.4.4 The choice of support spaces

One of the most important steps in the implementation of both a GME and GCE optimization problem concerns the choice of the support spaces defined for each of the unknown parameters by the researcher.
Concerning this issue, some "guiding principles" are suggested in researches in existing literature.
Firstly, it is worth mentioning that the choice of the intervals for the support space should be such that the "true" value of the parameter is within the bounds. Any vector of support points corresponding to a given parameter implicitly contains inequality restrictions represented by the largest and smallest support points.
However, as stressed by Lansink (1999), the chosen interval should be wide enough to allow for feasible solutions to the maximization problem.
Where we do not have knowledge about the unknown coefficients, the best choice is to specify, symmetric supports around zero and with "large" negative and positive bounds, as underlined by Shen and Perloff (2001).
Rezek and Campbell (2007) stated that in the absence of a compelling economic theory, these bounds are set wide enough to be non-binding. In this case, they are generally constructed symmetrically giving the parameters an expected value of zero.
Whenever prior information or economic theories can be called upon, the highest or lowest bounds can be specified to restrict the plausible values of a coefficient to be either non-positive or non-negative. Since the parameter estimates simply represent convex combinations of the support points, this specification guarantees the theoretically consistent sign on the parameters.
Concerning the number of the support points, Golan et al. (1996a) underlined that adding more points to the support of **Z** should decrease the variance of the associated point estimator in GME problems. These Authors constructed a sampling experiment based on 10000 Monte Carlo trials and a specified experimental design in order to check the impact of **M**. On the base of these limited results, it appears that the greatest improvement in precision comes from using M equal to 5 number of support points.

Regarding the choice of **V**, it is firstly important to note that the elements of **V** reflect the variation (e.g. variance, support) of the underlying errors, and the bounds may be functions of the T observations.
In actual fact, the unobservable disturbance vector, **e**, may represent one or more sources of noise in the observed system, including sample and no-sample errors in the data, randomness in the behaviour of the economic agents and specification or modelling errors, if $\mathbf{X\beta}$ is a convenient approximation of the underlying system.
The choice of the number of support points may be used to express or recover additional information about e_t (e.g. skewness or kurtosis).
If we assume the error distribution is symmetric and centred around 0, we can specify a symmetric support $v_{t1} = -v_{tJ}$ for each t where J is usually fixed at 3. Golan *et al*. (1996a; 2008) suggested that the choice of the support space for the

error should be made according to the 3-sigma rule, meaning that the error bounds should be set at three times the standard deviation from the origin.
The support space for the errors will be therefore on the form [-3stdev; 0; 3 stdev], where st.dev. is the empirical (from the data) standard deviation of the dependent variable.

In short, by re-parameterizing the unknown coefficients in terms of probability distributions on a finite and discrete support, the solution values of the parameters are consistent with the prior beliefs about the range of plausible parameter values and with the available sample. For this reason, as stressed by Lence and Miller (1998) the effective bounds placed on the parameter and sample spaces by the supports may be viewed as a virtue of GME. However, as stressed by the same Authors, if we have little or no information about the plausible values of the model parameters, the need to specify such supports may be viewed as the key drawback of the GME approach.
In addition to these statements, Lence and Miller (1998) found that the GME results are not sensitive to changes in the width of the error supports, and the changes in the parameter supports must be relatively large to have an impact on the parameter and input estimates.
It is usually recommended to perform a sensitivity analysis on the obtained estimates by making a moderate change in the support space specified for the unknown parameters. Golan et al. (2001) in the context of an AIDS for the Mexican meat demand, found that by making a moderate change in the support vectors while keeping the centre of the support unchanged, had negligible effects on the estimated coefficients and elasticities.
However, Paris and Caputo (2001) underlined that by carrying out a complete comparative static analysis of the GME estimator for the general linear model, it is possible to show that nothing can be said, "a priori" concerning the direct response of the estimates to changes in either parameter or error bounds. This research will not attempt to arbitrate on such matter. In fact, we believe that entropy is a useful and practical way of incorporating prior information whenever researchers are reasonably certain that parameters lie within a region.

2.4.5 Diagnostic measures

A simple way for evaluating the estimated coefficients is to compare the obtained values to the "a-priori" (from the theory) expectations in terms of sign and magnitude (Fraser, 2000; Nganou, 2004).
However, different tools were introduced for assessing the statistical validity of the estimated coefficients. These statistics are part of the output provided in SAS, LIMPDEP, SHAZAM and other software which includes the GME procedure.
One of the best known and applied evaluation measures is the normalized entropy (NE) measure that quantifies the relative informational content in the data. This measure is obtained as the ratio between the Shannon entropy measure calculated on each of the unknown parameters and the maximal value of the Shannon entropy function.

In actual fact, it is possible to calculate two different normalized entropy measures which quantify the information contained in the data, by using the same formulation.

For each variable *k,* the entropy measure is a continuous function from zero to ln (M). In order to make a comparison, these entropy measures are normalized to the zero-one interval.

The NE measure $S(\cdot)$ for the whole model is:

$$S(\hat{\mathbf{p}}) = \frac{-\sum_{k,m} \hat{p}_{km} \ln \hat{p}_{km}}{K \ln M} \quad [2.44]$$

where k=1,…,K refers to the variables while M refers to the number of support points defined for the support space of the unknown parameters.

The normalized entropy defined by the [2.44] lies in the interval ranging between 0 and 1, $S(\hat{\mathbf{p}}) \in [0,1]$ with $S(\hat{\mathbf{p}})$=0 and $S(\hat{\mathbf{p}})$=1 reflect the absence of uncertainty and complete uncertainty, respectively.

By using this information index it is also possible to establish if additional information or even restrictions in the data (expressed in the form of restrictions) produce a reduction of uncertainty and consequently a reduction in the basic uncertainty relative to the observed phenomenon. It is worth noting that this information measure is conditioned by the choice of Z.

It is interesting to note that, this measure also reflect the reduction of uncertainty achieved from the estimation problem.

The numerator of the [2.44] indicates the entropy related to the data information, while the denominator indicates the maximum level of uncertainty, which is the entropy level of the uniform distribution with M outcomes, for each of the K defined variables.

In order to assess the amount of information in each parameter, it is possible to use the NE measure as an information theory statistic reflecting the information in each of the variable k=1,2…K, which can be described as:

$$S\left(\hat{p}_k\right) = \frac{-\sum_{m} \hat{p}_{km} \ln \hat{p}_{km}}{\ln M} \quad [2.45]$$

This specific information measures reflect the relative contribution (for explaining the dependent variable) of each of the independent variables. The above normalized entropy measure can be connected to Fano’s inequality and the probability of errors (Golan, 2008).

As stressed by Tonini and Jongeneel (2008) it possible to calculate the NE measure, as described by the [2.45], separately for the signal and the noise parts of the model. In this way, they provide measures for the importance of the contribution of the parameters and each error term.

Concerning the NE measure, Fraser (2000) states that values near one mean that the solution is nearly uniform and that the data agrees with the prior information. For values near zero, the prior and the data reflect different information about the parameters and therefore the GME solution is non-uniform.
In the diagnostic and inference contexts, Mittelhammer and Cardell (1996) proposed asymptotic standard errors for the estimates and performed simple *t-tests*.
In addition to the diagnostic tools described above it is possible to use the overall degree of fit measures (R^2 and Adjusted R^2) for the estimated equations. This overall goodness of fit measure remains a useful summary statistic although it is said to be biased downward in GME cases, due to its use of out of sample information (Fraser, 2000).
Particularly, the R-square obtained from the GME case, as pointed out by Nganou (2004), will tend to be lower than the R-square obtained by using the OLS estimator.
The determination of standard errors of the obtained estimates will be a further development of this research, and in the near future we aim at obtaining parameter distributions for the GME and GCE estimates by using the bootstrap method as well.

2.5 The Use of Entropy Based Estimation Approaches in the Economic Context

Since their introduction, an important number of GME and GCE applications have appeared in the empirical economics literature and have been applied to different economic contexts and situations.
These estimation approaches have always been attractive, according to Heckelei, Mittelhammer and Jansson (2008) mainly for two reasons. On one hand, by using these methods it is possible to specify and to estimate under-determined models, which is a capability that is not provided by classical solution estimation methods. On the other hand, prior information on the unknown parameters can be included in a very simple way, making estimates potentially more efficient in a mean square error sense, or at least more "plausible" for model simulation, interpretation and analysis following estimation.
One of the most significant areas of application refers to balancing large raw data sets by using accounting identities and prior information to fill gaps and reconcile conflicting data sources.
These methods allow the researcher to set ranges for missing data values and provide a means for differentiating the reliability of various sources in the balancing process (Robinson, Cattaneo and El-Said, 2000; Britz and Wieck, 2002; Robilliard and Robinson, 2003).
In particular Robinson et al. (2000), used a flexible CE approach starting from inconsistent data estimated with error, for dealing with the problem of incorporating and reconciling the information from several sources and in order to estimate a SAM for Mozambique.

Robilliard and Robison (2003) presented an approach for reconciling household surveys and national account data. These authors studied the problem concerning how to use the information provided by the national account data in order to re-estimate the household weights used in the survey, so that the survey results were consistent with the aggregate data. In order to solve this problem statistically they took advantage of CE estimation criterion. They implemented their approach for Madagascar by underlining the power and flexibility of this approach in their results since it guarantees an efficient use of information from a variety of sources to reconcile data at different levels of aggregation in a consistent framework.

However, GME and GCE applications are not only used for data recovery and calibration issues, and have been employed in the attempt to solve traditional estimation problems better or analyze new problems (Golan, Judge and Perloff, 1996b; Zhang and Fan, 2001).

In short, any economic model characterized by a vector of M equations in K>M unknowns, is an underdetermined model that can be solved by using GME or GCE techniques.

Golan, Karp and Perloff (1996d) proposed the GME-Nash approach to estimate firm strategies consistent with the game theory and the underlying data generation process. In their research, they proposed the GME-Nash estimator which was compared to the Maximum Likelihood (ML)/ME estimator. The authors underlined that the proposed estimator allows for a greater flexibility than the ML-ME estimator; in particular they proved that it is more efficient in terms of mean square error, correlation and other measures of variances.

Whenever production and demand function parameters or the RCGE/CGE framework are called for, GME and GCE approaches appear to be among the most attractive estimation methods, because they allow the researcher to successfully deal with the frequent problems of incomplete and insufficient data which lead to the already known “ill-posed” estimation problems.

Lence and Miller (1998) used the GME approach as a feasible means for estimating production parameters when the activity-specific input allocations are not available. After re-parameterizing the production function parameters and the unknown disturbances, the maximum entropy framework may be used to recover estimates of the unknown parameter, error, and input values that are consistent with the available data and with the assumed production structure.

Lansink (1999) used the GME method to estimate a dual model of production based on panel data of Dutch cash crop farms over the period 1970-1992. In particular, the GME estimator allows the author to define a coherent system of input demand and output supply to be estimated for each farm in the sample, thus capturing technological heterogeneity. The estimation results are used to perform a cluster analysis to identify groups of farms with similar technologies.

Fraser (2000) used the GME and GCE methodologies in order to estimate a set of demand relationships, which are usually subject to a high degree of collinearity among the explanatory variables, for a UK meat demand data set. In his paper, the author makes some interesting comments about the potential of these estimation methods.

Zhang and Fan (2001) adapted the GME approach to empirically estimate multi-output production functions and input allocations for Chinese agriculture and for each Chinese province.

Arndt et al. (2002), in the context of a CGE model for Mozambique (Arndt, Cruz Jensen, Robinson and Tarp, 1998), provide an interesting application of the GME/GCE methods, in order to estimate the elasticity associated to aggregate CES and Constant Elasticity of Transformation (CET) functions, by using time series data and some “a priori” information on the parameters.

Using the GCE estimator, Balcombe et al. (2003) carried out an interesting study whose results supplied the estimates of the AIDS concerning consumption in Greece. In addition to this, they reported some interesting comments regarding these entropy based estimators made by both Classical and Bayesian statisticians. In actual fact, they underlined that for classical statisticians the inclusion of “subjective” information is a common source of contention. On the other hand, Bayesian statisticians are likely to object to the fact that entropy estimation does not construct what they might see as properly formulated posterior distributions. The same authors underlined that there is no need for “priors” to be highly informative within the entropy approach and however, good prior knowledge concerning support spaces sometimes exists. Moreover, from an entropic point of view, the Bayesian approach amounts to a rule (using Bayes theorem) that requires *ad hoc* choices for both the prior and likelihood functions, that together give a posterior distribution.

Nganou (2004) applied GME methodology to estimate Armington’s demand elasticity - representing the degree of substitution between domestic and imported goods - for some commodities in the context of a CGE model for the small South-African country of Lesotho.

Howitt and Msangi (2006) proposed GME estimates for disaggregate production functions regarding the Northern Mexico area. Specifically, these authors used values from a calibrated optimization model to define support spaces which were centred on values that were feasible solutions to the data constraints and consistent with prior parameter values.

Ferrari and Manca (2008) used the GME approach to estimate the parameters of a CES production function for the Italian region Sardinia by using a regional SAM extended to the environment (RESAM).

An interesting application is provided by Tonini and Jongeneel (2008), who used the GME approach for modelling the milk and beef supply for Hungary and Poland by reconciling sample and non-sample information. It is worth noting that in their research the authors introduced non sample information in the GME approach via microeconomic-theoretical constraints on parameters and constraints of the model originating from knowledge based on other economic and non-economic research. In the discussion of the results, they underlined that the final parameter estimates, obtained by taking advantage of an estimation process which used all the available information, were close to the original sample data, widely consistent with the economic theory and which complied with their expectations.

All of these interesting and successful applications of the GME and GCE estimators encouraged us to apply these entropy-based methodologies to our estimation problem too. In actual fact, as already mentioned, the GME and GCE philosophies allow us to translate an "ill-posed" problem, namely the obtainment of RCGE behavioural parameters by only using the information in the RSAM, into a "well-behaved" statistical problem with a unique solution.

Chapter 3

Functional Forms and Data

"...Social accounting is intended to classify, measure and present the transactions which take place over a period in an economic system in such a way that as far as possible they will accord with economic definitions and distinctions and as a result will be useful for economic analysis especially as it relates to practical economic policy"
[R.Stone]

3.1 The Role of Production and Demand Functions in the RCGE Framework

In the implementation of a RCGE model, the issue of the specification of a particular functional form, which illustrates producer and consumer behaviour, recovers a key role.
As stressed by Annabi et al. (2006) the choice of the functional forms is critical, given that the estimation results may be sensitive to different model specifications. Therefore, functional forms appear to be influential in RCGE and CGE model performance.
In actual fact, within the framework of the RCGE modelling, economists need to consider elasticity corresponding to each behavioural function used[1]. Most of the existing applied studies have often used functional forms which do not allow for flexibility in the values of elasticities parameters and above all in the value of elasticity of substitution parameters.
In order to overcome the problems concerning lack and scarcity of data, which are among the most important obstacles to the computation of CGE models both at national and sub-national level, applied CGE modellers often use behavioural functional forms whose unknown parameters could be entirely calibrated from the RSAM (or from the SAM for the national CGE model).
Furthermore, in the case of sub-national level, like a region, these difficulties are much more probable. However, this circumstance, thanks to the entropy based methods, does not prevent us from obtaining the behavioural parameter values, in order to introduce them in the RCGE model implemented for Sardinia.
In this research, the selection among alternative competing functional forms concerning the structure of production and demand spheres is based on a careful

[1] As already mentioned the majority of the modellers often prefer to borrow the values of elasticities from the literature, since the direct estimation requires the availability of the statistical data specific to each sector and each country.

examination of both theoretical and empirical aspects by taking into consideration also the type of data available.
In order to decide the functional forms which will go to specify the relationships in the production and demand environments, the first necessary step consist in describing the general structure of the production and demand sphere in the studied context. It is worth noting that the two following paragraphs will be focused on the theoretical and empirical aspects (referred to the Sardinian economy structure) of the chosen functional forms, which are the CES function and the Working-Leser function for the production and household demand spheres, respectively.

Production sphere in a RCGE model context is commonly described by using a *multi-stage production process*, with each stage containing a different set of factors in order to allow for different elasticity of substitution in each level and between different set of factors.[2] Moreover, each producer is assumed to maximize profits, defined as the difference between revenue earned and the cost of factors and intermediate inputs. Following the standard literature on RCGE and CGE models (Lofgren, Harris and Robinson, 2002; Partridge and Rickman, 2008) the top-down structure of this *production tree* can be explained as follows.
At the top level, regional output is modelled as a function of value added and intermediate consumption. The Leontief production functional form is usually chosen to represent the production of regional output with fixed proportions of composite value added and composite intermediate inputs.
At the second level, the composite primary factors generally enter the production process which allows different factor substitution degree to be specified in order to consider different factor substitution degrees. In particular, primary factors, labour, land and capital, join together to generate value added. This relationship is generally described through CD or CES functional forms, where the former implicitly specifies unitary factor substitution elasticities while the latter is a more general case which gives a value of substitution elasticity which is not necessarily equal to one. At the same level, intermediate goods from different countries/regions join together to form composite intermediate goods (which will be included in the first level of production). To model this relationship between the two categories of intermediate inputs, the CES function is commonly applied.[3]
The structure of a RCGE production process like the one described above, as stressed by Partridge and Rickman (2008) and other authors, distinguishes this

[2] As Vargas, Scheiner, Tembo and Marcouiller (1999) argued, these models follow the neoclassical economic theory, where the factor demand depend on both output and relative factor prices and as a consequence of this, the RCGE models (but obviously the CGE models too) do not represent factor demands as linear function of output.
[3] A third level in the production process may represent substitution among labour skills within the overall labour input, among classes of land within the overall land input, or types of capital inputs within the overall classification of capital (Vargas et al., 1999). A common procedure is to consider the CES form of production which allows substitution elasticity to differ among industries. It is important to note that RSAM does not always show subcategories of primary inputs to model this type of relationship, so the most important relationships are those expressed by the two levels mentioned.

family of models from a simple Input-Output model (I/O), since demand factors do not linearly only depend on the output, but also on the sensibility of the output towards factor costs and towards the prices of intermediate goods.
While the underlying assumption in regional I/O models and in the more general quantitative representation of a regional economy provided by a RSAM is the Leontief technology, production specification in RCGE models is made in a neoclassical theory framework. In almost all cases, production technology is portrayed as a compromise between these two views, with intermediate goods and services based on the Leontief technology – i.e. linearly depending on output – and sensitive to prices, factor demands that depend non-linearly on both output and factor prices.
In particular, in order to specify the top level relationship between value added and intermediate consumption the Leontief functional form is often chosen. In fact, this functional form is characterized by a substitution elasticity equal to zero and the unknown parameters can be entirely calibrated from the RSAM. Usually the Leontief functional form is the default in this relationship. However, as stressed by Lofgren et al. (2002) the CES alternative may be preferable if particular evidence suggests that available techniques allow the aggregate mix between value-added and intermediate inputs to vary.
On the other hand, in order to model the second level relationship, CD functions are often specified. Although a greater degree of substitution between factors is introduced, this function specifies a unitary and constant substitution elasticity and the unknown parameters, namely the efficiency and distribution parameters, are provided in the RSAM.
Whenever more flexible functional forms, such as the CES or the Transcendental Logarithmic (TRANSLOG) production functions, are introduced in the computation process, the related unknown parameters are seldom directly estimated empirically since researchers prefer to “borrow” them from the estimates available in literature.
By using a CES production function, the only unknown parameters has been represented up to now by the substitution elasticity since the efficiency and the share parameters have been calibrated from the RSAM. To this respect, an important innovative aspect of our approach is that all the unknown parameters are obtained as a result of a simultaneous estimation process.
In this context, Salem (2001) estimated the elasticity of substitution of the CES production function, between capital and labour, for the case of Tunisia by using statistical data coming from I/O Tables for the period 1985-1994 for 14 sectors which produce tradable goods. In order to be able to estimate these elasticity for each economic sector and considering the chosen functional form, he applied the method of Non-Linear Least Squares (NLS) to the software GAMS[4].

[4] The obtained values of elasticities of substitution for each sector explain well the existence of a CES relation between the two factors of production, capital and labour. Salem found that the majority of elasticity are lower than 1 except for the textiles, clothing and leather sector (since the elasticity is equal to 1) and for the oil /gas sector (2.575) and the Transport and telecommunication sector (1.556).

In particular, the family of the CES production functions has been widely used in CGE models since it allows a considerable degree of flexibility in model specification. According to the existing literature (Kouparitsas, 2001) (Lofgren et al. 2002), our choice has been to model all the relationships characterizing the economy of Sardinia by CES production functions.

Concerning demand and specifically the household demand analysis for consumption needs, it is firstly important to stress that households wish to attain a certain consumption level at the lowest possible expenditure according to the neoclassical logic. One of the most important issues studied by CGE modellers, regarding demand is demand elasticity. In fact, these elasticity values have important implications for the CGE models.

As stressed by Sánchez (2004) trade policy reform will induce changes in domestic relative prices. This will result in a reallocation of resources, which in turn will affect the returns to production factors and the amount of factor earnings, mostly labour income, transferred to households. This income effect, as underlined by the same author, will enhance or constrain the capability of households to afford their basic consumption.

In general, the knowledge of income or expenditure elasticity – since the expenditure is usually assumed as a proxy of the income in the demand analysis – reflects the responsiveness of household demand towards an income change (or towards expenditure changes induced by an income change).

A functional form used to model the consumption block, as stressed by De Boer and Missaglia (2006) is the LES for which the Engel curves are straight lines. In their research, the same authors proposed the Indirect Addilog System (IAS) in order to model the demand sphere of the Palestinian CGE model for the year 1998. By using this function they were able to estimate income elasticity for the CGE model studied by allowing for non-straight Engel curves, inferior commodities and elastic demand.

In our research, we based the demand system specification on the Working (1943)-Leser (1963) functional form, by taking into consideration the economic meaning and structure of the RSAM, which can be considered as a one-period "snapshot" for the structure of a particular economy.

3.2 Modelling the Production Sphere: the CES function

3.2.1 Main properties, characteristics and the economic theory guidelines

Several economic studies have used the CD functional form in order to model the production function and find relationships among economic variables. As stated by Kadiyala (1972) this choice has often been made since the parameters of this function can be easily determined and it is quite consistent with most economic data. However, the assumptions on which the CD function is based[5] are not always consistent with the specific economic context analyzed[6].

More specifically, the assumption of unitary elasticity of substitution between factors, that characterizes the CD function, has not been checked and confirmed by empirical studies.
In 1961 Arrow, Chenery Minhas and Solow (ACMS[7]), firstly argued that economic analysis, based on the CD assumptions, often led to conclusions that were unduly restrictive. In particular, in their seminal work beginning with an empirical observation[8] these authors pointed out that the elasticity of substitution is always not equal to one. These empirical findings encouraged these authors to find a mathematical function which contemporaneously had: i) the property of homogeneity; ii) a constant elasticity of substitution between two factors (i.e. capital and labour) and iii) the possibility of different elasticities of substitution for different industries, sectors or countries.
On the base of this evidence and considering the above-mentioned requirements, Arrow et al. specified a general production function with these properties and which is described by the following expression:

$$Y = \alpha\left[\delta X_1^{-\rho} + (1-\delta) X_2^{-\rho}\right]^{-\frac{1}{\rho}} \qquad [3.1]$$

where Y is the output, X_1 and X_2 represent the two inputs.
The parameter α represents the *efficiency parameter*: it serves as an indicator of the state of technology and it plays the same role as the coefficient α in the CD function. More precisely, Arrow et al (1961) refers to the α parameter as a neutral efficiency parameter since a change in the parameter α changes the output for any given set of inputs in the same proportion.

[5] Kadiyala (1972) argued that the CD production function has some serious drawbacks: i) it requires all inputs to be positively employed and ii) it has a unitary elasticity of substitution.
[6] CGE modellers have often used the CD functional form to model the production system, since the parameters of this function can easily calibrated from the RSAM.
[7] The ensemble of the first letters of the surname of each proposer of this functional form, namely ACMS, has become an alternative name of this production function.
[8] Before exploring the possible functional forms of the "their" production function they tested, with a simple regression analysis, the relations among the value added in thousands of US dollars (V), labour input in man-years (L) and money wage rate (W, obtained by the ratio between total labour cost and L), described by the two following equations:

$$\frac{V}{L} = c + dW + \eta$$

$$\log\frac{V}{L} = \log a + b\log W + \varepsilon$$

From the regression analysis, they found that in 20 out of 24 industries, over 85 per cent of the variation in labour productivity is explained by variation in wage rates alone (Arrow et al. 1961).
In a linear relationship like the above, Arrow et al. (1961) underlined that the elasticity of the dependent variable (i.e. the ratio V/L) with respect to W is constant and equal to b. At the same time, this production function will have a constant elasticity of substitution equal to b.

The parameter δ represents the *share (distribution) parameter* between the two input factors and it represents the relative factor distribution in the production process.
The parameter ρ, as defined by Arrow et al. (1961) is a transformation of the elasticity of substitution σ and will be called the *substitution parameter*. In fact, it is the substitution parameter which determines the value of the (constant) elasticity of substitution σ[9]. For this reason, the substitution parameter plays a crucial role in the CES production function.
The CES production function is characterized by a homogeneity of degree one[10]. To demonstrate this, let us replace X_1 and X_2 by jX_1 and jX_2, where j is a constant that multiplies each input factor; the output will change from Y to jY as explained by the following expression:

$$\begin{aligned} Y &= \alpha \cdot \left[\delta \left(jX_1 \right)^{-\rho} + \left(1-\delta\right)\left(jX_2\right)^{-\rho} \right]^{-\frac{1}{\rho}} \Rightarrow \\ &= \alpha \cdot \left\{ j^{-\rho} \left[\delta X_1^{-\rho} + \left(1-\delta\right) X_2^{-\rho} \right] \right\}^{-1/\rho} = \left(j^{-\rho} \right)^{-1/\rho} Y = jY \end{aligned} \qquad [3.2]$$

For this reason, the CES function has constant returns to scale, like all linearly homogeneous production functions and possesses average products and marginal products that are homogeneous of degree zero, due to the properties of function homogeneity of degree one, in the variables X_1 and X_2.[11]
The family of production functions described by the CES functional form includes all those functions which have a constant elasticity of substitution for all values of the ratio of two inputs X_1 and X_2.

---

[9] As better specified later the elasticity of substitution between the two factors, σ, is obtained as: σ=1/(1+ρ)

[10] In general a function is said to be homogeneous of degree *r*, if multiplication of each of its independent variables by a constant *j* will alter the value of the function by the proportions j^r, that is, if:

$$f\left(jx_1,...,jx_n\right) = j^r f\left(x_1,...,x_n\right)$$

In general *j*, can take any value. However, in order for the above equation to make sense, $\left(jx_1,...,jx_n\right)$, must not lie outside the domain of the function f. For this reason, in economic applications the constant *j* is usually taken to be positive, as most economic variables don't admit negative values. (Chiang, 1984)

[11] According to Giusti (1994), the marginal productivity of the two factor X_1 and X_2 are:

$$\frac{\partial Y}{\partial X_1} = \delta \alpha^{-\rho} \left(\frac{Y}{X_1} \right)^{1+\rho} ; \frac{\partial Y}{\partial X_2} = \left(1-\delta\right) \alpha^{-\rho} \left(\frac{Y}{X_2} \right)^{1+\rho}$$

The dX_1 and dX_2 ratio is the negative of the slope of an isoquant; that is, it is a measure of the marginal rate of technical substitution of 1 for 2. Thus, the slope of isoquants (with X_1 plotted vertically and X_2 plotted horizontally) is:

$$\frac{dX_1}{dX_2} = -\frac{\partial Y / \partial X_2}{\partial Y / \partial X_1} = -\frac{1-\delta}{\delta} \left(\frac{X_1}{X_2} \right)^{1+\rho}$$

The economic theory is a guide for the plausible values of the three parameters which characterize the CES production function.
The parameter α reflects the productive efficiency and therefore cannot be negative. More specifically for a fixed combination of the input factors, a variation in α gives product changes in the same proportions. For this reason it has been named efficiency parameter, as stated by Giusti (1994).
The parameter δ, ranges between 0 and 1 since it describes the relative contribution of each input in the production process.
The admissible values of the ρ parameter are strictly related to the admissible values for the elasticity of substitution σ. A more detailed analysis concerning the relation between ρ and σ can be seen in the next paragraph[12].

Although the CES function allows the economists to overcome some assumptions belonging to the CD or the Leontief production functions which are considered too restrictive, this production function has been criticized for several reasons.
One of the most debated issues concerns its non linearity (Gebreselasie, 2008). In fact, the CES function cannot be analytically linearized[13], even in its logarithmic form. Therefore estimating functional parameters for the CES function must necessarily include non-linear fitting techniques which could be usually quite complicated both in the case of two inputs and in the case of *n*-inputs .[14]

3.2.2 The role of the ρ parameter and the elasticity of substitution

The degree of substitutability between the input factors of a production function is an essential concept within the production theory. Hicks (1932) was the first to introduce and discuss a dimensionless measure of substitutability of input factors, the so-called *elasticity of substitution* for a two-factor production.
More specifically, the Hicks elasticity of substitution, that is a pure number (Estrin and Laider, 1995) measures the percentage change in the factor ratio divided by the percentage change in the marginal rate of technical substitution (MRTS)[15], with output being held fixed (Laureti, 2006).

---

[12] It will be interesting to note that the CES function, for particular values of the elasticity of substitution includes the Leontief and the CD functions as its special cases. Also for these reasons the CES function is often used in the CGE models.
[13] Both the logarithmic of the CD production function and the translog form are linear; this makes these forms easy to estimate employing standard econometric packages. Also, the two input forms of these, as underlined by Hoff (2004) are easily extended to *n* inputs, and the resulting forms, also being linear, are easy to apply.
[14] In order to solve these kind of problems it has been proposed firstly by Kmenta (1967) a linearization of the two-input CES function, employing a Taylor approximation, the result of which is a restricted form of the general translog function. However the linearization of the CES function represents only an approximation and it is further only applicable for a certain range of CES parameters.
[15] The formal derivation of the marginal rate of technical substitution can be expressed as follows:
$MRTS_{ij} = \dfrac{\partial f / \partial X_i}{\partial f / \partial X_j} = -\dfrac{dX_j}{dX_i}$ and $MRTS_{ji} = \dfrac{\partial f / \partial X_j}{\partial f / \partial X_i} = \dfrac{1}{MRTS_{ij}}$
Considering a production function with two inputs x_1 and x_2 and along an isoquant consider a particular change in which only factor 1 and factor 2 change, and the change is such that output

According to Varian (1992), this is a relatively natural measure of curvature of an isoquant. It explains how the ratio of factor inputs changes as the slope of the isoquant, measured by the MTRS, changes. If a small change in slope gives a large change in the factor input ratio, the isoquant is relatively flat which means that the elasticity of substitution is large.
By using the logarithmic derivative, that is the common practice in economics in order to obtain this measure, it is possible to re-write the elasticity of substitution as (Giusti, 1994):

$$\sigma_{12} = \left[\frac{d \ln\left(x_2/x_1\right)}{d \ln\left|\frac{\partial f / \partial x_1}{\partial f / \partial x_2}\right|} \right]_{Y=\bar{Y}} = \sigma_{21} \qquad [3.3]$$

Mishra (2008) in his interesting review on production functions underlines that the elasticity of substitution for the CES function is constant along and across the isoquants.
Moreover, as already stressed above, it is the substitution parameter that determines the value of the elasticity of substitution.
Applying the [3.3] in the CES production function, the elasticity of substitution is described as[16]:

$$\sigma = \frac{d \ln\left(x_2/x_1\right)}{d \ln\left|MRTS\right|} = \frac{1}{\rho+1} \qquad [3.4]$$

The value of σ can be anywhere between 0 and ∞; the larger the value of σ, the greater the degree of substitution between the two inputs. The two extreme cases can be described as follows. The former, that is when σ is equal to 0, describes a situation in which the two inputs are used in a fixed proportion, as complements to each other. The latter, with σ equal to infinite, is where the two inputs are perfect substitutes for each other. (Chiang, 1984).

---

remains constant. (That is, d x_1 and d x_2 adjust along an isoquant). Since output remains constant, we can write:
$0 = \frac{\partial f}{\partial x_1} dx_1 + \frac{\partial f}{\partial x_2} dx_2$, which can be re-written as $\frac{\partial f}{\partial x_2} dx_2 = -\frac{\partial f}{\partial x_1} dx_1$, which can be solved for $\frac{dx_2}{dx_1} = -\frac{\partial f/\partial x_1}{\partial f/\partial x_2} = -MRTS_{12}$

16 According to Varian (1992), since the marginal rate of tecnical substitution can be espressed as: $TRS = -\left(\frac{x_1}{x_2}\right)^{-(\rho+1)}$ so that $\frac{x_2}{x_1} = \left|TRS\right|^{\frac{1}{\rho+1}}$ by taking logs it becomes $\ln\frac{x_2}{x_1} = \frac{1}{\rho+1}\ln\left|TRS\right|$, so that the expression [3.4] returns.

To analyze this relation in more detail, and referring to a CES production function with two inputs, namely X_1 and X_2, it is firstly necessary to introduce the expressions of the marginal products of X_1 and X_2, respectively.
As shown by the equation [3.4] the CES production function takes on a variety of shapes depending on the value of the parameter ρ. In such sense, the CES function contains several other well-known production functions as special cases, depending on the value of the parameter ρ.
On the basis of the admissible values of the elasticity of substitution ($0 \leq \sigma \leq +\infty$) the lowest admissible value for ρ is -1. This implies an infinite elasticity of substitution, $\sigma_{\rho=-1} = \frac{1}{\rho+1} = \frac{1}{-1+1} = \frac{1}{0} = \infty$, and therefore the two inputs are perfect substitutes and the isoquants are straight-lines.
For values of ρ between -1 and 0, we have elasticities of substitution greater than unity.
The case of ρ=0 yields a unitary elasticity of substitution which is characteristic of the Cobb-Douglas function. In this way, the CD function is a special case of the (linearly homogeneous) CES function. This does not appear immediately from the equation [3.1] since when ρ=0 the CES function is undefined. Nevertheless, it is possible to demonstrate that for values of ρ that approach to zero the isoquants of the CES production function look very much like the isoquants of the Cobb-Douglas production function.
For values of ρ greater than zero we have elasticity of substitution less than 1. As ρ approaches to +∞, a CES isoquant looks like an isoquant associated with the Leontief technology and therefore there is not possibility of substitution between the two inputs, that are combined in fixed proportions in the production process.

3.2.3 The CES functions in the Sardinian RCGE context

Regarding the Sardinian RCGE model, which represents the applied economic framework of this research, the production system is described with a two-stage process. As usual, *at the first level* value added and intermediate inputs join together to obtain the regional output. *At the second stage*, value added is assumed to be a function of capital and labour incomes.
Each level, from the highest to the lowest, is modelled by using the CES functional form. Although the selected models seem to differ from the usual specifications applied in empirical literature, the specification introduced could be considered more general and at the same time more flexible, since the CES production function includes the Leontief (generally used to model the top-level) and the CD specifications (generally used to model the second level relationships) as special cases.
These remarks ensure that the results of the analysis will absolutely agree the existing theory even if a more general specification is adopted. Precisely, whenever the substitution elasticity should be equal to zero in a CES function the Leontief functional form returns. At the same time, for values of elasticity of substitution equal to one the CES production function reduces to the CD functional forms.

For our estimation problem, we will refer to the logarithmic form of the CES production function, which can described by the following expression (Salem, 2001; Mishra, 2008):

$$\ln Y = \ln \alpha - \frac{1}{\rho} \ln\left[\delta X_1^{-\rho} + (1-\delta) X_2^{-\rho}\right] \qquad [3.5]$$

and where the parameters have the same role and meaning of the traditional form of the CES introduced by [3.1].

The theoretical specification of the CES production function applied to the production process of the Sardinian RCGE model follows the original formulation introduced by Arrow et al (1961) which is expressed by the [3.1] expressed in logarithmic form like the [3.5].

Concerning the *top-level relationship*, in which value added and intermediate consumptions are combined to obtain regional output, the specified functional form is expressed as:

$$\ln Y_i = \ln \alpha_i - \frac{1}{\rho_i} \ln\left[\delta_i VA_i^{-\rho_i} + (1-\delta_i) IC_i^{-\rho_i}\right] + \varepsilon_i \qquad [3.6]$$

where Y_i, VA_i and IC_i are the regional output, the value added and the intermediate consumptions of the branch *i*, respectively[17].

At *the second level*, the CES function that combines capital and labour incomes to obtain value added, is expressed as:

$$\ln VA_i = \ln \alpha_i - \frac{1}{\rho_i} \ln\left[\delta_i CI_i^{-\rho_i} + (1-\delta_i) LI_i^{-\rho_i}\right] + \varepsilon_i \qquad [3.7]$$

where VA_i, CI_i and LI_i are the value added, the capital incomes and the labour incomes of the branch *i*, respectively[18].

It should be stressed that the proposed approach allows for more flexibility not only concerning the parameterization of the behavioural functions, but it also introduces more flexibility on the degree of substitution between factors both in the first and in the second level of the production process.

[17] Details on the definition of the variables will be discussed in the paragraph 3.4.

[18] It is worth noting that estimates have been carried out for both the two production stages, separately. To avoid repetitions, results will be described in detail in Chapter 4 for the top-level relationship. However, in the Appendix 1 we well provide the estimates for the second-stage relationship (namely concerning value added as function of labour and capital incomes) as the final results of the proposed three-step estimation approach.

3.3 Modelling the Demand Side: Working Leser Functional Form

3.3.1 Main characteristics and properties

The approach we adopted to model the demand side is included in the Engel curve framework, traditionally used to model consumer behaviour (Deaton and Muellbauer, 1980a, 1980b).
The Engel curve represents the relationship between commodity expenditure and income and its shape plays an important role in consumer demand analysis.
Leser (1963) in his seminal work underlined that the problem of finding the most appropriate form of an Engel function is a well-known problem in econometrics, "…but as yet no solution appears to have found general acceptance".
In his study, Leser proposed the following functional form that relates the value of the expenditure share to the logarithm of total expenditure, as follows:

$$w_t = \alpha_t + \beta_t \ln(X) + \varepsilon_t \qquad [3.8]$$

where, w_t ($w_t \geq 0$) is the relative quota or budget share of the expenditure for the "good" *t* over the total expenditure and X indicates the total expenditure. The α_t and β_t represent unknown parameters to be estimated and ε_t represents the error term. This formulation is generally known as the Working – Leser functional form since it was firstly proposed by Working in the 1943 and then further developed by Leser in 1963. This form, as underlined by Castaldo and Reilly (2007) is consistent with household utility maximization.
Since the seminal papers by Working (1943) and Leser (1963), many studies have found and validated a log-linear relation between income and commodity expenditure share. This relation is also motivated by the popular AIDS models of Deaton and Muellbauer (1980a, 1980b) and Jorgenson, Lau, and Stoker (1980) which represent the natural extension of the Working Leser model when price information are required to introduce in the model.

The model is based on the assumptions that: i) all budget share sum to unity; ii) the adding-up restrictions, which requires that: $\sum_t \alpha_t = 1$ and $\sum_t \beta_t = 0$.
The Working Leser model, as underlined by De Mello, Pack and Sinclair (2002) rests upon a particular class of preferences, the PIGLOG class (which means logarithm of prices independent generalized linearity), which are represented with a cost or expenditure function defining the minimum expenditure required for gaining a specific level of utility at given prices. On this issue, Muellbauer (1976) showed that the PIGLOG class of preferences allows for an exact aggregation over consumers without imposing identical preferences.

3.3.2. The expenditure elasticity

One of the most important issues of a demand relationship system study concerns, as already mentioned, the obtainment of income-expenditure elasticity, where expenditure is assumed to be a proxy of the income.
The expenditure elasticity is a measure of the responsiveness of demand of a particular commodities (or a group of similar commodities) to changes in the total expenditure. As noted by Akbay Boz and Chern (2007) this measure is used to approximate corresponding unconditional income elasticities.
To be precise, the expenditure elasticity shows how the quantity purchased changes, that is how sensitive it is, in response to a change in the consumer's expenditure.
The analytical formulation of the expenditure elasticity can be described as the ratio between percentage change in quantity demanded and percentage change in expenditure.
For the Working Leser functional form, using the definition of elasticity, the expression of the expenditure elasticity can be derived as follows (De Mello et al., 2002; Chern, Ishibashi, Taniguchi and Tokoyama, 2003; Castaldo and Reilly, 2007):

$$\eta_t = \left(\beta_t + w_t\right)\frac{1}{w_t} = \frac{\beta_t}{w_t} + 1 \qquad [3.9]$$

The economic meaning of the expenditure elasticity is that if the percent change in the quantity demanded is greater than the percent change in consumer expenditure, the demand is said to be expenditure elastic, or responsive to changes in consumer expenditure. On the other hand, if the percent change in the quantity demanded is less than the percent change in consumer expenditure, the demand is said to be expenditure inelastic, or not responsive to changes in consumer expenditure.

3.3.3 The Working Leser Model for Sardinia

The study of the demand relationships and the consumer expenditure patterns can give an indication of the demand tendency and growth in a particular economy.
Modelling the household consumption sphere in a macro-economic context, like the RSAM and consequently like the Sardinian RCGE context, means studying the consumption sphere within an extremely aggregated information context.
Since a RSAM (or equivalently a SAM) is based on data which refers to a single year observation, prices related to the demanded quantity are assumed to be constant. As established by the Theory of Consumption (Deaton and Muellbauer, 1980b) the demand functions are derived from the maximization problem of the consumer who make the demanded quantity of a good or service dependent on the income level of the individual and on prices.
However, when prices are constant, as stressed by Beneito (2003) one obtains the specification corresponding to the Engel curve, where the demand of goods is

dependent only on the income of the individual. In such a case, the only condition to be met is the adding-up restriction since all the others are derived from the consideration of prices.
Bearing this in mind, our choice was to relate the value of the expenditure shares to the logarithm of total expenditure following the original statistical analysis of budget shares by Working (1943) and Leser (1963).
The peculiarity of the RSAM at our disposal allows us to carry out household consumption analysis by considering different groups of households. Thanks to this classification it will be possible to emphasize the different expenditure shares and the different expenditure patterns which refer to the various income levels.
The analysis that will be carried out, whose results will be shown in Chapter 5, has different interesting aspects since it is aimed to study the demand elasticity of the household towards different production branches.
The general expression of the Working-Leser function specified for the Sardinian household demand system is the following:

$$w_{ji} = \alpha_{ji} + \beta_{ji} \ln(X_j) + \varepsilon_{ji} \qquad [3.10]$$

where j refers to household income group studied while i refers to the economic production branch. Further details on this models will be specified in Chapter 5.

3.4 Data: Regional SAM for Sardinia

3.4.1 Some notes on the SAM

The SAM, was firstly introduced by Richard Stone in 1962, as a database able to arrange flows in a coherent and economically meaningful way, which are expressed in terms of value and characterize an economic system in accountancy (Ferrari, 1999).
In this way, a SAM represents an example of such a consistent database which can be defined as a numerical representation of the economic cycle. Sánchez (2004) underlines that the SAM represents an economy-wide data set that can be used to feed CGE model equations.
The SAM as stressed by Robinson (2003) provides a framework for data organization, multiplier analysis, macro models, and multi-sectoral models such as CGE models.
With particular reference to the RCGE models it is important to remember that the most crucial data requirement in developing this kind of model is represented by a recent and "good" RSAM for the particular region studied.
The RSAM represents a snapshot of an economy in a particular year, to which the RCGE is calibrated[19]. Therefore regional SAM databases and regional CGE

[19] Consequently, the base run of the model replicates the initial economic equilibrium captured in the SAM.

modelling frameworks share the same economic features and can be interpreted as different ways of representing the national economy.

It is important to note that the RSAM is constructed in such a way that it contains the I/O table introduced by Leontief, namely in the sub-matrix describing intermediate consumption. In this way, it is possible to state that the RSAM refers to the technological structure of the I/O Table.
As an extension of an I/O flow table, a RSAM is also a useful system in exploring the macroeconomic implications of policy changes. The RSAM framework allows more complete economic analysis, resulting from any changes in policies in comparison with a fixed coefficient I/O model.
As stressed by Pyatt and Round (1985) a further objective of the SAM is to provide the statistical basis for the creation of a plausible model. For this purpose a SAM approach integrates the distributional dimension within the system of national accounts in a way that reflects the interrelationship between employment, distribution of income and the structure of production[20]. In particular the SAM usually focuses on the distribution of income through disaggregation of household sector income and expenditure accounts together with disaggregation of production, factors etc.
According to these interpretations, and in alignment with Stone's theory (1962), the SAM, both at national and regional level, represents contemporarily: i) a *macro-accounting structure* ii) a *matrix* with all the properties of a true matrix and iii) an *economic model*.
As an *accounting structure* a SAM is the result of an aggregation process concerning economic agents and institutions, no longer dividable, such as households, firms, government and rest of the world, each of them make transactions with the others in the economic system. The object of these transactions concerns flows in terms of value, such as the distribution of income (among different types of households) or of value added (among the different factors of production) (Thorbecke, 2000).
In short, the SAM can be seen as a set of different subjects aggregated on the basis of their role in the economic system, and depending on i) the reason for which they carry out the transaction, and ii) their position in the social structure or sectoral location (Ferrari, 1998).
As a *matrix* and more precisely as a square matrix a SAM represents a mathematical structure in which the functional and sectoral units record their transactions and where it is important to bear in mind the convention that entries are read as receipts for the row account in which they are located and as expenditure for the corresponding column account[21].
To sum up, the payments (expenditures) are listed in columns and the receipts are recorded in rows. As the sum of all expenditures from a given account (or

[20] In other words, the SAM summarizes succinctly the interdependence between production activities, factor shares, household income distribution, balance of payments, capital accounts, etc. for the economy as a whole at a point in time.
[21] On this issue Torbercke (2000) underlines that within the SAM "each transactor or account has its own row and column".

subaccount) must equal the total sum of receipts or income for the corresponding account and therefore row sums must equal the column sums of the corresponding account[22].

At the same, the RSAM framework (Ferrari, 1999), as well as the SAM framework, must satisfy two basic rules (Pyatt, 1999):

i) For every row there must be a corresponding column and the system is only complete if the corresponding row and column totals are identical;

ii) Every entry is a receipt when it is read in its row context and expenditure when it is read in its column. The description of SAMs as single entry accounts derives from this rule.

The SAM is also an *economic model* since it includes the three main functions of an economic system, - production, consumption and accumulation – under the Keynesian economic framework.

According to these three interpretations of the SAM, it is worth emphasizing that the SAM, as well as the RSAM, represent flexible tools for modelling the structure of an economic system at a level of disaggregation which is coherent and agrees with the desired analysis, which is obviously subject to available data. In fact, a SAM captures transactions in an economy, regarding both income and expenditure, like double-entry accounting, but it contains much more information than a normal macro aggregate. A SAM, even if it is similar to a standard I/O model which systematically include production relationships, it has the added advantage of capturing income distribution and consumption relationships within an economy in an internally consistent way[23].

3.4.2 The RSAM for Sardinia

As already mentioned, the main aim of this research is to estimate production and demand function parameters in a "self-contained" approach by only using the information contained in a SAM and in particular in a RSAM.

The dataset under which the estimates will be carried out is represented by the RSAM of Sardinia for the year 2001 (Ferrari, Garau and Lecca, 2007).

The RSAM for Sardinia at our disposal was originally disaggregated into 23 production branches.

An interesting feature of the available RSAM concerning the household income allocations is represented by the subdivision of the Sardinian household into 6

[22] For example, the total income of a given institution (say a specific socioeconomic household group) must equal exactly the total expenditures of that same institution.. Hence, as stressed again by Torbercke (2000), analysts interested in understanding how the structure of production influences the income distribution can obtain useful insights by studying the SAM.

[23] Bearing these rules in mind it could be seen that SAM describes basic transformations in the economy. Production in the economy uses the intermediate materials and the primary factors of production. These factor supplies are contributed by the institutions, who, in turn, receive factor payment as value added. In addition to value added, institutions might get income from other sources such as transfers. (Sinha, 2008)

groups, which gives us a better understanding of the primary income allocation process[24].
In the aim of estimation and in order to increase the power of our estimates we aggregated the production sphere of the original RSAM into six production branches according to the NACE (*Classification of Economic Activities in the European Community*) classification concerning economic activities.
By doing so, the RSAM used for the estimation process has a production sphere which includes the following economic activities:

i) Agriculture, animal husbandry and fishing;
ii) Energy and Mining products;
iii) Industrial products;
iv) Construction
v) Market services;
vi) Non Market services;

For the sake of clarity, it is important noting that each one of these six production branches includes a different number of branches, as summarized in Table 3.1. In detail:

i) The *Agriculture, animal husbandry and fishing* branch is composed of the two following branches:
- Agriculture, hunting and forestry;
- Fishing and related services;

ii) The *Energy and Mining products* branch includes the following branches:
- Mining and quarrying;
- Manufacture of non metallic mineral products;

iii) The *Industrial product* branch includes the following branches:
- Manufacture of food products, beverages and tobacco;
- Manufacture of leather and leather products;
- Manufacture of pulp, paper and paper products; publishing and printing;
- Manufacture of coke, refined petroleum products; Chemical and pharmaceutical products;
- Manufacture of basic metals and fabricated metal products;
- Manufacture of machinery and equipment; manufacture of electrical and optical equipment; Manufacture of transport equipment;
- Manufacture of wood, rubber and plastic and related products; other manufactured products;
- Electricity, gas and water supply;

[24] The classification of households is of crucial importance in a SAM analysis. In the SNA 1993, the importance of household classification has been explicitly discussed. This document states that "Conclusions regarding (changes in) inequality, and perhaps even poverty, may have to be based on subgroup averages, and thus depend very much on how the population has been subdivided.

iv) The *Construction* branch is composed of one branch and therefore the name and the numerical amount of this aggregate branch and those of the related disaggregated branch coincide;

v) The *Market service* branch is composed of the following branches:
- Wholesale and retail trade; repair of motor vehicles, motorcycles etc…;
- Hotels and restaurants;
- Transport, storage and communication;
- Financial intermediation;
- Real estate, renting, research and development and other business activities;

vi) The Non Market service branch is composed of the branches:
- Public administration and defence; compulsory social security;
- Education;
- Health and social work;
- Other community, social and personal service activities;

Table 3.1 – Passage from the 23-branch RSAM to the 6-branch RSAM for Sardinia (2001)

Sardinian RSAM	Number of branches included
Agriculture, Animal Husbandry and Fishing	2
Energy and Mining Products	2
Industrial Products	9
Construction	1
Market Services	5
Non Market Services	4
Total	23

Source: our elaboration on the RSAM for Sardinia (2001)

It is important to note that the branches included in each of the aggregate branch represent at the same time the number of the observations under which our estimation processes are based.

Regarding household classification, the same aggregation concerning the production branches was maintained.
Moreover, household income distribution is divided into six household groups by considering different income levels and namely:

i) *Household income group A*: disposable annual income lower than 9,300 Euro;
ii) *Household income group B*: disposable annual income from 9300 to 12400 Euro;
iii) *Household income group C*: disposable annual income from 12,400 to 15,500 Euro;

iv) *Household income group D*: disposable annual income from 15,500 to 24,800 Euro;
v) *Household income group E*: disposable annual income from 24,800 to 31,000 Euro;
vi) *Household income group F*: disposable annual income higher than 31,000 Euro;

The aggregated RSAM for Sardinia is shown in Figure 3.2, where the values reported are in Millions of Euro.
The Sardinian RSAM matrix offers a level of intermediate consumption distinguished for the six specified branches.
Intermediate consumption is defined as the value of products, which are transformed or used up as inputs to a process of further production.
For each branch the amount of intermediate consumption referred to the same itself branch or to the other branches can be seen by the cells from the (1,1) to the (6,6). Specifically each column, within these cells, identifies the intermediate consumption characterizing each of the six branches.
Value added is specified in Income from Capital (CapInc), Income from Labour (LabInc) plus social contributions (SocContr). This specification means that value added at this form represents the amount remaining for distribution to the primary factors and equals the total value of factor incomes generated by production.

Table 3.2 shows the amount, in millions of Euro, of the value added and the intermediate consumptions for each branch included in the Sardinian RSAM for the year 2001.

Table 3.2 Value added and intermediate consumption for each branch (in Millions of Euro)

Sardinian RSAM	Value added	Intermediate consumption
Agriculture, Animal Husbandry and Fishing	1039.654	574.604
Energy and Mining Products	341.701	607.230
Industrial Products	2882.636	7430.501
Construction	1473.257	2039.906
Market Services	11948.257	7739.105
Non Market Services	6799.783	3357.496

Source: our elaboration on the RSAM for Sardinia (2001)

Intermediate consumption and value added are the macro-economic aggregates used to model the production system of Sardinia region.
In the estimation process, described in the following chapter, we will consider value added and intermediate consumption as the two inputs which combine to produce the total regional output for each production branch.

Rows 10 and 11 record net taxes and subsidies on production, respectively, while rows 12 records the Value Added Tax (VAT).
As already mentioned, households are detailed in six group rested on their annual disposable income.

Table 3.3 shows the different expenditure amounts toward the six branches for each household group.

Table 3.3 Household consumption expenditure (values in millions of Euro) towards the production branches

Aggregate RSAM	**Hh A**	**Hh B**	**Hh C**	**Hh D**	**Hh E**	**Hh F**
Agriculture, animal husbandry and fishing	65	89	51	113	24	35
Energy and mining products	7	17	10	20	9	9
Industrial products	778	1244	824	1919	477	656
Construction	5	5	4	7	2	3
Market Services	1253	1959	1488	3460	962	1242
Non Market Services	189	427	188	414	114	132
Total expenditure	*2297*	*3741*	*2565*	*5933*	*1588*	*2077*

HhA: households pertaining to the income group A;
HhB: households pertaining to the income group B;
HhC: households pertaining to the income group C;
HhD: households pertaining to the income group D;
HhE: households pertaining to the income group E;
HhF: households pertaining to the income group F;
Source: our elaboration on the RSAM for Sardinia (2001)

Finally, rows from 22 to 24 record tax on imports and imports from Italy and from the Rest of the World.

Table 3.4 RSAM for Sardinia (year 2001) – Aggregation into six branches

		1	2	3	4	5	6	7	8	9	10	11	12	13A	14B	15C	16D	17E	18F	19	20	21	22	23	24	TOTAL
1	AGRAHF	174	1	504	1	46	12							65	89	51	113	24	35		1	1		779	6	1901
2	ENEMIP	3	199	1363	520	14	10							7	17	10	20	9	9			35		354	31	2601
3	INDUP	235	198	3638	621	1866	927							778	1244	824	1919	477	656		9	2439		3570	1923	21323
4	CONSTR	0,47	5	144	398	203	89							5	5	4	7	2	3			2898		19	0,4	3782
5	MARKSER	158	200	1713	489	5325	1384							1253	1959	1488	3460	962	1242		150	1009		418	629	21839
6	NMARKSER	5	4	69	13	285	935							189	427	188	414	114	132		7502	22		320	1	10619
7	LabInc	263	136	966	523	2984	3444																			8316
8	SocContr	57	58	362	183	819	1352																			2831
9	CapInc	719	148	1555	768	8145	2003																			13338
10	Ind Tax	6	5	1054	42	199	215																			1521
11	ProdContr	-103	-4	-76		-192	-2																			-377
12	Vat	15	6	591	193	812	87																			1704
13	Household A							719	276	442				10						289	551					2287
14	Household B							746	297	752					34					979	1799					4608
15	Household C							491	155	1111						8				389	915					3068
16	Household D							3309	1125	1801							13			566	1234					8049
17	Household E							1201	392	672								4		186	422					2877
18	Household F							1849	585	1356									3	173	406					4373
19	Firms									7158				43	61	43	80	26	32		912					8354
20	Publ Adm									46	1521	-377	1704	509	543	485	412	309	294	3291			124			8860
21	Sav													-572	230	-31	1610	949	1968	2482	-5042			1968	2841	6404
22	M tax	3	1	119		0,4																				124
23	M IT	201	540	5660	21	851	154																			7427
24	M RoW	163	1104	3662	12	481	10																			5432
	TOTAL	1901	2601	21323	3782	21839	10619	8316	2831	13338	1521	-377	1704	2287	4608	3068	8049	2877	4373	8354	8860	6404	124	7427	5432	151261

Source: Our elaboration on RSAM for Sardinia (Ferrari, Garau and Lecca, 2007).
Legend: AGRAHF: Agriculture, Animal Husbandry and Fishing; ENEMIP:Energy and Mining Products; INDUP: Industrial Products; CONSTR: Construction; MARKSER: Market Services; NMARKSER: Non Market Services;

Chapter 4

Three step Entropy Approach for the Estimation of the CES Production Function

4.1 The Specified Functional Form and the Estimation Strategy

The theoretical specification of the CES production function applied to the production process of the Sardinian RCGE model follows the original formulation introduced by Arrow et al (1961) which is described by the [3.1].
For the estimation procedures, the logarithmic transformation of the CES production function is used, although its non-linearity in the logarithmic transformation, has been one of the most important obstacles to its econometrically application. It is worth to emphasize that in the GME and GCE estimation approaches the non-linearization of the CES production function does not cause any problems.
Concerning the relationship in which value added and intermediate consumptions are combined to obtain total regional output, which describes the highest level of the multi-stage production system defined for Sardinia, the specified functional form is expressed as:

$$\ln Y_i = \ln \alpha_i - \frac{1}{\rho_i} \ln\left[\delta_i VA_i^{-\rho_i} + \left(1-\delta_i\right) IC_i^{-\rho_i}\right] + \varepsilon_i \qquad [4.1]$$

where Y_i, VA_i and IC_i are the total output, the value added and the intermediate consumptions of the branch *i*, respectively[1].
As usual, the unknown parameters in the expression [4.1] are α_i which represents the efficiency parameters describing "the state of technology", δ_i which is the distribution parameter and ρ_i which is the substitution parameter; while ε_i represents the error component.
After having estimated the parameter ρ_i it will be possible to obtain the elasticity of substitution which characterizes each level and each branch, by using the already mentioned relation $\sigma_i = \left(1/(1+\rho_i)\right)$.
The index *i* refers to the branches for which estimates will be carried out. Specifically, in the models to be estimated, data aggregation leads to the following six branches (see Chapter 3 for the aggregation criteria):

[1] Details on the definition of the variables are available in chapter 3.

i) *Agriculture, animal husbandry and fishing;*
ii) *Energy and mining products ;*
iii) *Industrial products ;*
iv) *Construction ;*
v) *Market services;*
vi) *Non Market services;*

The base year for the model is 2001, which corresponds to the last year for which a detailed RSAM for Sardinia is available.
Although the selected model seems to differ from the usual specifications applied in empirical literature for this kind of relationship, the specification introduced could be considered more general and at the same time more flexible, since the CES production function includes the Leontief (generally used to model this top-level relationship) and the CD specifications as special cases.
These remarks ensure that the results of the analysis will absolutely agree the existing theory even if a more general specification is adopted. Precisely, whenever the substitution elasticity should be equal to zero in a CES function the Leontief functional form returns. At the same time, for values of elasticity of substitution equal to one the CES production function reduces to the CD functional forms.

However, the most essential aspect of the proposed approach is that a full estimation of the CES unknown parameters is carried out, by only using the macro information contained in the RSAM. On this issue, it is worth noting that, while the underlying assumption in regional I/O models and in the more general quantitative representation of a regional economy provided by a RSAM is that of Leontief technology, production specification in RCGE models is made in neoclassical theory framework.
Therefore, on one hand our estimation approach aims to recover the value of unknown parameters by a process based on the RSAM information only. On the other hand, since we insert in the RCGE framework, where different and flexible functional forms are generally used to describe the economic agents' behaviour, there was the need to make the approach more general, that is to say not only strictly related to the Leontief technology under which a RSAM is based. In order to accomplish these requirements, we will introduce a *three-step estimation process*.
By following this strategy, in the *first step* any "a priori" information about the order of sizes of the parameter values and any prior knowledge about the parameter bounds will not be introduced thus allowing the data to speak. Then, *in the second step* additional information will be gradually introduced, by taking advantage of the results obtained in the first step together with the economic requirements. Finally, *the third step* follows a GCE philosophy on the basis of which prior information is introduced both in terms of parameter support bounds and in terms of an estimated prior probability distribution.

Further details on the suggested estimation strategy together with some possible further improvements, will be discussed in the paragraph 4.5, after the obtained results will have been illustrated.

4.2 The First Estimation Step: GME Approach

4.2.1 The GME re-parameterization and optimization problem

In order to estimate the parameters of the equation [4.1] by using the GME method firstly we need to re-parameterize all the unknown parameters and the error components and express all of them in terms of probabilities. Clearly, this procedure was carried out for each branch.
To provide the reader with a feeling of how GME operates, here the estimation procedure is reported only for the *Agriculture, Animal Husbandry and Fishing*[2] branch, in order to avoid repetition.
For this branch, the CES production, before the GME re-parameterization, can be expressed as:

$$\ln Y_{AGRAHF} = \ln \alpha_{AGRAHF} - \frac{1}{\rho_{AGRAHF}} \ln\left[\delta_{AGRAHF} VA_{AGRAHF}^{-\rho_{AGRAHF}} + \left(1-\delta_{AGRAHF}\right) IC_{AGRAHF}^{-\rho_{AGRAHF}}\right] + \varepsilon_{AGRAHF} \quad [4.2]$$

Regarding GME estimation purposes, as stated by Golan (2008), each of the unknown parameters is viewed as the mean value of some well defined random variable rather than searching for their point estimates. In the same manner, the unobserved error component is also viewed as another set of unknowns, and similar to the signal component, each ε_i is constructed as the mean value of some random variable *v*. Clearly, the main objective is to estimate the unknown parameters with minimal distributional assumptions.

Concerning our estimation problem, firstly we have to express all the parameters and the error term in terms of proper probabilities.
For example, in order to transform the parameter α, we will start by choosing a set of discrete points, $\mathbf{a_i} = (a_{i1},...,a_{is},...,a_{iS})'$ where *i* refers to the specific analyzed branch while *s* represents the dimension S≥2 of the support space.
We will consider the dimension of the support space for all the parameters and for all the branches equal to 5 (i.e. for the efficiency parameter α, S=5).
After defining the support space dimension, a vector of corresponding unknown weights is also introduced as follows: $\mathbf{p_i} = \left(p_{i1},...,p_{is},...,p_{iS}\right)'$ so that $\sum_{s=1}^{S} p_{i,s} = 1$ and $p_{is} \geq 0$, which are the requirements that a proper probability distribution must

[2] From here on the Agricultural, animal husbandry and fishing branch is named AGRAHF.

satisfy. At this point we are able to re-write the parameter α_i, before the estimation process, as $\sum_{s=1}^{S} a_{is} p_{is} = \alpha_i$.

The same procedure was used for the other unknown parameters and namely for the distribution parameter δ and for the substitution parameter ρ.

Obviously, this procedure was also carried out for all the branches analyzed.

More precisely:

i) a set of 5 discrete points, $\mathbf{t}_i = (t_{i1}, ..., t_{im}, ..., t_{iM})'$, was introduced for the distribution parameter δ, where *i* refers to the specific analyzed branch and *m* refers to the dimension of the support space (M≥2) which is equal to 5. The vector of corresponding unknown weights is also introduced as follows: $\mathbf{b}_i = \left(b_{i1}, ..., b_{i,m}, ..., b_{iM}\right)'$ under the already mentioned normalization constraints, so that $\sum_{m=1}^{M} b_{i,m} = 1$ and $b_{i,m} \geq 0$. In this way, it is possible to re-write the distribution parameter δ as $\sum_{m=1}^{M} t_{im} b_{im} = \delta_i$;

ii) the substitution parameter ρ has been re-parameterized by introducing a set of 5 discrete points, $\mathbf{z}_i = (z_{i1}, ..., z_{id}, ..., z_{iD})'$ where *i* refers to the specific analyzed branch and *d* refers to the dimension D≥2 and equal to D=5 of the support space. The vector of corresponding unknown weights is also introduced as follows: $\mathbf{q}_i = \left(q_{i1}, ..., q_{i,d}, ..., q_{iD}\right)'$ under the already known normalization constraints, that $\sum_{d=1}^{D} q_{i,d} = 1$ and $q_{i,m} \geq 0$. In such a way, it is possible to re-write the substitution parameter ρ as $\sum_{d=1}^{D} z_{id} q_{id} = \rho_i$.

Similarly, in order to re-parameterize the error component ε_i we went on to obtain a transformation by specifying a vector of H≥2 discrete points $\mathbf{v}_i = (v_{i1}, ..., v_{ih}, ..., v_{iH})'$, distributed uniformly around zero, where *i* refers to the analyzed branch.

A vector of proper unknown weights $\mathbf{w_i} = \left(w_{i1}, ..., w_{ih}, ..., w_{iH}\right)'$ is associated to the support space, so that $\sum_{h=1}^{H} v_{ih} w_{ih} = \varepsilon_i$. As recommended by Golan et al. (1996a) the number of the points of the error support space was fixed at H=3, while the three sigma rule (where sigma is empirical – from the data - standard deviation of the dependent variable) and the symmetry around zero was used for the support space of the error term[3].

[3] It is worth noting that for the Construction branch we used a symmetric support [-1;1] according to Ferrari and Manca (2008).

As stated by Golan et al. (2001) and Nganou (2004) there is no need to assume any subjective information on the distribution of the probabilities with the GME. The re-parameterized coefficients lead to the following expression for the CES production function, which will enter in the GME optimization problem:

$$\ln Y_i = \ln\left(\sum_{s=1}^{S} a_{is} p_{is}\right) - \frac{1}{\left(\sum_{d=1}^{D} z_{id} q_{id}\right)} \ln\left[\left(\sum_{m=1}^{M} t_{im} b_{im}\right) VA_i^{\left(-\sum_{d=1}^{D} z_{id} q_{id}\right)} + \left(1 - \left(\sum_{m=1}^{M} t_{im} b_{im}\right)\right) IC_i^{\left(-\sum_{d=1}^{D} z_{id} q_{id}\right)}\right] + \sum_{h=1}^{H} v_{ih} w_{ih} \qquad [4.3]$$

The related optimization problem is based on the following objective function:

$$\underset{\{p,q,b,w\}}{Max}\; H(\mathbf{p},\mathbf{q},\mathbf{b},\mathbf{w}) = -\mathbf{p}'\ln(\mathbf{p}) - \mathbf{q}'\ln(\mathbf{q}) - \mathbf{b}'\ln\mathbf{b} - \mathbf{w}'\ln\mathbf{w} \qquad [4.4]$$

subject to the data constraint (described by the [4.3]) and the adding-up constraints (for each *i*) for the probabilities attached to the parameters and the noise terms, that are:

$$\sum_{s=1}^{S} p_{is} = \sum_{m=1}^{M} b_{im} = \sum_{d=1}^{D} q_{id} = 1; \qquad [4.5]$$
$$\sum_{h} w_{ih} = 1;$$

The solution to this maximization problem is unique. Forming the Lagrangean and solving for the first-order conditions yields the optimal solution for the proper probabilities. The point estimates for the unknown parameters are recovered from the estimated probabilities and the specified support points, as follows: $\widehat{\alpha}_i = \sum_s a_{is}\widehat{p}_{is}; \hat{\delta}_i = \sum_{m=1}^{M} t_{im}\hat{b}_{im}; \widehat{\rho}_i = \sum_d z_{id}\widehat{q}_{id}$.

4.2.2 The choice of the support spaces

Bearing in mind that estimation results may be sensitive to support specification[4] and also considering the type of data available, the *first step* of the estimation strategy used, will specify wide and equal supports for all the parameters[5].
In order to evaluate how the various support spaces affect the resulting estimates, five different models for each branch were specified.
Each model differs from the others in the support space bounds specified for the unknown parameters[6]. More precisely, as already proved by Golan et al (2001), a moderately large change in these support vectors was made while keeping the

[4] As underlined by Fraser (2000) the estimated parameter values could change sign and increase in magnitude.
[5] To perform the estimation GAMS, using the MINOS/PATHNLP solvers, was employed.
[6] For example the model named GME 1 specifies a wide support, between -100 and 100 both for alpha, delta and rho and centred on zero. The model GME2 defines the support bounds between -50 and 50 and it is always centred on zero.

centre of the support unchanged, in order to evaluate the effects on the estimated coefficients both in terms of their estimated sign and in terms of magnitude.
At the same time, the magnitude of the obtained coefficient values could represent a good starting point to define more informative supports for the parameters, which will be discussed later and will be the aim of the second step of the estimation strategy used.
The choice to include also negative values for all the three parameters, even though the economic theory could give us some "guiding principles", has been done in this first estimation step in agreement with the GME theory, that suggest to define a symmetric support centred on zero, so that if the data does not add information the estimated values converge to zero. Therefore, by examining the results shown in Tables 4.1-4.6 it is possible to evaluate to what extent the obtained estimates are sensitive and robust to the change of the supports. As underlined by Nganou (2004), this procedure represents a simple way to evaluate and to interpret the estimated coefficients, since it is possible to compare the obtained parameter values with the expectations of the economic theory in terms of their signs and magnitude.
By examining the results concerning the *Agriculture, Animal Husbandry and Fishing* branch (Table 4.1), the first point to note is the relative robustness of the estimates for the parameters α and ρ. In particular, both the α and the ρ parameters have the right sign in agreement with the economic theory. The estimated values for the δ parameter are robust and correct in terms of sign, but they also seem to be relatively variable in terms of their magnitude. Although they are of the correct magnitude in almost all the models, in the runs 3 and 4 the estimated values are greater than one and this is not acceptable according to the theory.

Table 4.1 – Sensitivity Analysis of GME Estimates of the CES function for Agriculture, Animal Husbandry and Fishing

	GME 1	GME 2	GME 3	GME 4	GME 5
α	1.969 (0.9997)	1.837 (0.9991)	1.499 (0.9965)	1.252 (0.9902)	1.052 (0.9722)
δ	0.857 (0.9999)	0.919 (0.9997)	1.088 (0.9982)	1.148 (0.9918)	0.951 (0.99773)
ρ	0.0702 (0.9999)	0.025 (0.9999)	0.018 (0.9999)	0.08 (0.9999)	-0.0002 (0.9999)
Entropy value	7.025	7.023	7.016	6.99	6.92
NE	0.9999	0.9996	0.9982	0.9940	0.9831

The normalized entropy measure for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-4.27;4.27]
GME2: [-50;50] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-4.27;4.27]
GME3: [-20;20] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-4.27;4.27]
GME4: [-10;10] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-4.27;4.27]
GME5: [-5;5] for α, δ, and ρ; [-3stdev;3stdev] for the error;
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.2 - Sensitivity Analysis of GME Estimates of the CES function for Energy and Mining products

	GME 1	GME 2	GME 3	GME 4	GME 5
α	13.520 (0.9886)	13.768 (0.9521)	10.031 (0.8340)	6.924 (0.6613)	4.22 (0.4266)
δ	2.402 (0.9996)	2.722 (0.9981)	2.052 (0.9934)	1.181 (0.9913)	1.603 (0.9107)
ρ	0.177 (0.9999)	0.712 (0.9999)	0.018 (0.9992)	1.271 (0.9899)	-0.083 (0.9998)
Entropy value	7.006	6.619	7.016	6.99	5.58
NE	0.9961	0.9834	0.9422	0.88089	0.7791

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.56;0.56]
GME2: [-50;50] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.56;0.56]
GME3: [-20;20] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.56;0.56]
GME4: [-10;10] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.56;0.56]
GME5: [-5;5] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.56;0.56]
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.3 - Sensitivity Analysis of GME Estimates of the CES function for Industrial products

	GME 1	GME 2	GME 3	GME 4	GME 5
α	3.831 (0.9990)	5.099 (0.9935)	2.696 (0.9887)	3.588 (0.9176)	2.561 (0.8263)
δ	-0.012 (1.000)	0.354 (0.9999)	-1.408 (0.9969)	0.037 (0.9999)	-0.223 (0.9988)
ρ	-1.667 (0.9998)	-0.027 (0.9999)	-2.07 (0.9933)	0.847 (0.9955)	1.649 (0.9306)
Entropy value	14.259	14.280	14.231	14.122	13.852
NE	0.9996	0.9978	0.9930	0.9710	0.9186

*The normalized entropy measures for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-2.59;2.59]
GME2: [-50;50] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-2.59;2.59]
GME3: [-20;20] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-2.59;2.59]
GME4: [-10;10] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-2.59;2.59]
GME5: [-5;5] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-2.59;2.59]
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.4 - Sensitivity Analysis of GME Estimates of the CES function for Construction

	GME 1	GME 2	GME 3	GME 4	GME 5
α	1.448 (0.9998)	1.447 (0.9994)	1.441 (0.9968)	1.423 (0.9873)	1.353 (0.9537)
δ	-0.726 (0.9999)	-0.725 (0.9999)	-0.719 (0.9992)	-0.701 (0.9969)	-0.638 (0.9898)
ρ	0.151 (0.9999)	0.151 (0.9999)	0.148 (0.9999)	0.139 (0.9999)	0.111 (0.9997)
Entropy value	5.927	5.926	5.920	5.901	5.828
NE	0.9999	0.9998	0.9986	0.9947	0.9811

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-1;1] for the error
GME2: [-50;50] for α, δ, and ρ; [-1;1] for the error
GME3: [-20;20] for α, δ, and ρ; [-1;1] for the error
GME4: [-10;10] for α, δ, and ρ; [-1;1] for the error
GME5: [-5;5] for α, δ, and ρ; [-1;1] for the error
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.5 - Sensitivity Analysis of GME Estimates of the CES function for Market services

	GME 1	GME 2	GME 3	GME 4	GME 5
α	2.357 (0.9996)	2.354 (0.9986)	2.348 (0.9914)	2.262 (0.9678)	2.031 (0.8934)
δ	0.241 (0.9999)	0.307 (0.9999)	0.432 (0.9997)	0.476 (0.9986)	0.581 (0.9916)
ρ	-2.596 (0.9996)	-1.613 (0.9993)	0.027 (0.9999)	0.042 (0.9999)	0.050 (0.9999)
Entropy value	10.319	10.316	10.302	10.260	10.107
NE	0.9997	0.9993	0.9970	0.9888	0.9616

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-1.5;1.5]
GME2: [-50;50] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-1.5;1.5]
GME3: [-20;20] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-1.5;1.5]
GME4: [-10;10] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-1.5;1.5]
GME5: [-5;5] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-1.5;1.5]
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.6 - Sensitivity Analysis of GME Estimates of the CES function for Non Market services

	GME 1	GME 2	GME 3	GME 4	GME 5
α	5.842 (0.9978)	2.549 (0.9983)	2.069 (0.9933)	2.048 (0.9737)	1.905 (0.9067)
δ	0.670 (0.9999)	0.905 (0.9998)	0.680 (0.9993)	0.701 (0.9969)	0.747 (0.9861)
ρ	0.00007 (1.000)	1.474 (0.9995)	-0.008 (0.9999)	0.030 (0.9999)	0.051 (0.9999)
Entropy value	9.216	9.217	9.196	9.160	9.022
Normalized Entropy Ratio	0.9997	0.9992	0.9975	0.9902	0.9642

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
GME1: [-100;100] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.77;0.77]
GME2: [-50;50] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.77;0.77]
GME3: [-20;20] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.77;0.77]
GME4: [-10;10] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.77;0.77]
GME5: [-5;5] for α, δ, and ρ; [-3stdev;3stdev] for the error, here we used [-0.77;0.77]
Source: our elaboration on the RSAM for Sardinia (2001)

The first point to note in Table 4.2, concerning *Energy and Mining products* is that α increases significantly as the support is widened, obviously this may have an impact on the other estimated values. In this respect, it is important to note that the estimates for δ even though they appear robust in terms of the estimated signs, they are always greater than one. The estimated values of the substitution parameter ρ present both admissible sign and magnitude, but they appear to be relatively variable to the support specification.
With regards to the *Industrial product* branch, the estimates for α are correct both in sign and the magnitude, although they seem to be quite variable to the support specification. The situation for the obtained values concerning the δ and the ρ is more difficult since the estimates appear to be rather unstable both in sign and magnitude in the different models.
On examining Table 4.4, the estimates derived for the *Construction* branch do not appear to be sensitive to the choice of the support values for all the parameters. The only “weak point” is represented by the sign of the share parameter δ, since it appears negative compared to the economic theory.
The results for the *Market service* branch are presented in Table 4.5 that shows the robustness of the α and δ parameters both in their magnitude and sign. On the contrary, the obtained estimates for the substitution parameter ρ seem to be quite sensible to the support choice even if they present the correct sign.
Finally, Table 4.6 shows the estimates for the *Non Market service* branch. Also in this case the estimates for all three parameters are correct both in sign and magnitude, even if they appear relatively sensitive to the support specification.

In short, it is possible to state that the obtained results are almost entirely coherent with the "guidelines" taken from the economic theory of the CES production function. Moreover, these estimates appear to be quite robust both in terms of sign and magnitude to the support specification even if they are relatively variable for some branches.
It is worth mentioning that in our determination of the final estimates which emerge from this *first estimation step*, shown in Table 4.7, we preferred those with results in keeping with the economic expectation and the widest support space. It is important to underline that in this case the economic theory could be considered a "diagnostic tool" in order to evaluate the validity of the GME estimation method, since any "a priori" information was not introduced in the estimation process.
This choice is in agreement with our logic, which does not take into consideration any "a priori" information and does not bias the estimates with a too tight support space.
On this issue, Golan et al. (1996a) underline that wide bounds may be used without extreme risks or consequences if our knowledge is minimal and we want to ensure that the support space contains the true values of the unknown parameters.

Table 4.7 – GME Estimates of the CES function (first estimation step)

	α	δ	ρ
Agriculture, animal husbandry and fishing[A]	1.969 (0.9997)	0.857 (0.9997)	0.0702 (0.9997)
Energy and mining products[B]	13.520 (0.9886)	2.402 (0.9996)	0.177 (0.9999)
Industrial products[C]	5.099 (0.9935)	0.354 (0.9999)	-0.027 (0.9999)
Construction[D]	1.448 (0.9998)	-0.726 (0.9999)	0.151 (0.9999)
Market services[E]	2.348 (0.9914)	0.432 (0.9997)	0.027 (0.9999)
Non Market services[F]	5.842 (0.9978)	0.670 (0.9999)	0.00007 (1.0000)

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
A: [-100;100] for α, δ, and ρ; [-4.27; 4.27] for the error
B: [-100;100] for α, δ, and ρ; [-0.56;0.56] for the error
C: [-50;50] for α, δ, and ρ; [-2.59;2.59] for the error
D: [-100;100] for α, δ, and ρ; [-1;1] for the error
E: [-20;20] for α, δ, and ρ; [-1.5;1.5] for the error
F: [-100;100] for α, δ, and ρ; [-0.77;0.77] for the error
Source: our elaboration on the RSAM for Sardinia (2001)

As shown in Table 4.7 the obtained results are almost entirely coherent with the expectations taken from the economic theory. This is completely true for the

Agriculture, animal husbandry and fishing branch, the *Industrial product*, the *Market service* and the *Non Market service* branches. For these production branches the CES estimated parameters range between the "admissible" economic values. However, the magnitude of the estimated parameters could be improved by specifying "aimed" support spaces for the unknown parameters of each branch. At this point of the research, several aspects would be worth noting regarding the choice of support values. In particular, in the light of results obtained a *second step* of the estimation strategy is implemented, by gradually adding some "a priori" information about the unknown parameters.

4.3 The Second Estimation Step: a Bounded GME (B-GME) Approach

4.3.1 The reasons underlying a B-GME approach

In the previous paragraph, GME estimates of the CES parameters describing the relationship between value added and intermediate consumption of the Sardinia's production system were carried out. This could be considered a "good" result since the estimation process was based on a very small amount of data. Therefore it is worth emphasizing the potentiality of the GME method since it enables us to model the production functions and to obtain the estimates in a situation characterized by minimal (or even negative) degrees of freedom. In fact, as stressed by several authors such as Tonini and Jongeneel (2008), this is only one of the main advantage of the GME estimators.
However, as shown in Table 4.7 some estimated coefficients do not agree with the expectation in terms of magnitude or sign suggested by the economic production theory.
These motivations represent the "fulcrum" which encouraged and motivated us to introduce a second estimation step, which was called B-GME. In particular, we maintained the methodological aspects of the estimator introduced by Golan et al (1996a) but, at the same time, a restriction on the "*a priori complete uncertainty*" was carried out. In fact, according to the GME methodological assumptions introduced in Chapter 2 it is clear that this method does not require "a priori" information either concerning the bounds of the unknown parameters or the probability distribution specified on the data, as stressed by Mittelhammer et al. (2000). On this subject, Golan (2008) stated that the GME minimizes the joint entropy distance between the data and the state of complete uncertainty (described, as well known, by a uniform distribution).
The proposed estimation step aims to gradually introduce some "a priori information" in the estimation process without "*jeopardizing*" the classical assumptions of the GME approach.
To be precise and referring to the econometric problem in question, a more informative supports on the efficiency and the distribution parameters were introduced, starting from the evidence of the previous step and some available "a priori" information.

By observing the results in Table 4.1- 4.6 we decided to introduce a restriction on the efficiency support spaces while maintaining the underlying uniform probability distribution. In particular, for the definition of the support the CES economic theory together with the information extracted from the RSAM were used.

Regarding the CES economic theory, the α parameter has the same meaning of the α parameter in the CD functional form. Specifically this parameter indicates the "state of technology" characterizing the production process and is called the *neutral efficiency parameter* . So, for its meaning and economic nature this parameter cannot be negative. This information was the first piece of information used to define the support space and specifically to establish the lowest bound of the support space.

Concerning the RSAM "aid" it is possible to determine from the matrix a calibrated value of the efficiency parameter which could guide us in the definition of the estimated expected size and magnitude of the α, as better explained in the following paragraph where the available information and the support specification will be formally introduced.

The second point to note concerns the support specification for the share parameter δ for which a support space ranging between 0 and 1 was introduced according to the CES economic theory.

Concerning the substitution parameter ρ the situation is quite different because a greater variability was found, above all in the estimated sign of the parameters, when the support space was changed. For this reason, we maintained a complete uncertainty on the support space bounds in this estimation step as well. This choice was made by considering that the support space for the other two parameters was "under control" in this second estimation step. Therefore, the sensitivity analysis carried out on the substitution parameter actually achieved a double objective. On one hand, it could be considered a "diagnostic tool" to understand how much information the estimated values of the substitution parameters add to the estimation process. On the other hand, by taking advantage of the sensitivity analysis, the support space which added most information to the estimated model was included in the third estimation step, *ceteris paribus* (the economic theory principles and the peculiarity of the studied context).

4.3.2 The available a priori information

In the definition of support regarding the efficiency parameter α, we took advantage of the calibrated efficiency measure from the available RSAM.

This value allowed us to get information on the degree of efficiency and the "state of the technology" concerning the production process, prior to the estimation procedure, on the basis of the underlying economic theory characterizing the SAM and without considering the underlying production functional form chosen for the RCGE studied framework (which in this study is represented by CES functional forms). In order to obtain this measure, characterizing every single branch, we used the macroeconomic information available within the RSAM. In particular, the amount of value added and intermediate consumption of each branch - which represented the inputs of our production function- were used and we carried out a calculation of the ratio of the two economic quantities. According to the National Accounting theory, the results obtained represent the calibrated degree of efficiency of a branch[7].

$$\alpha_i^{Calibrated} = \frac{\text{Value added}_i}{\text{Intermediate Consumption}_i} \qquad [4.6]$$

where, the index *i (i=1,…,6)* refers to the studied branches. The table including the single branch values (Table 4.8) can be seen below.

Table 4.8 - Calibrated efficiency parameters from the RSAM

	Calibrated α
Agriculture, animal husbandry and fishing	2.54
Energy and mining products	0.68
Industrial products	0.47
Construction	0.72
Market services	1.44
Non market services	3.09

Source: our elaboration on the RSAM for Sardinia (2001)

The "a priori" information regarding the distribution parameter δ was obtained by following a quite similar procedure to the one mentioned above for α.
More specifically, from the RSAM it is possible to determine a calibrated share parameter by calculating the ratio between the value of each input and the total amount of the inputs.

[7] It is important to note that the determined measure refers to the framework underlying the construction of the RSAM. Specifically concerning the production side and therefore the structure of an I/O Table, it is assumed that the two inputs enter the production process in fixed proportions; in this way no substitution degree is allowed for them. The production structure chosen in this study refers to a CES production function. This functional form, as stressed in chapter 3 allows for different substitution levels. Therefore, if the estimated substitution is different from zero, we can plausibly assume that the efficiency parameters will be different from those calibrated from the RSAM. This evidence represents another reason for underlining the importance of the estimation process and the estimated parameter values within the computation process of a RCGE model.

Table 4.9 shows the obtained values for the top-level relationship between the regional output and the two inputs, namely added value and intermediate consumption.

Table 4.9 - Calibrated distribution parameter from the RSAM

	Calibrated δ
Agriculture, animal husbandry and fishing	0.7028
Energy and mining products	0.3925
Industrial products	0.3074
Construction	0.4190
Market services	0.5744
Non Market Services	0.6727

Source: our elaboration on the RSAM for Sardinia (2001)

4.3.3 The choice of the support spaces and the B-GME estimation results

On the basis of the calibrated values for the efficiency and distribution parameters a focused and "informative" support was defined.
Concerning the efficiency parameter α, the lowest bound was fixed at zero as required by the economic theory[8]. The highest bound was chosen by considering the obtained calibrated efficiency parameter. Bearing this in mind, our choice was, at this stage, to fix an highest bound for all the branches which was wide enough to not bind the estimation process, taking also into consideration the peculiarity of each of the studied branch and the aggregation criteria[9]. By doing so, in the support space we clearly included the values calibrated from the RSAM.
Our choice concerning the definition of the support spaces of the *distribution parameter δ*, depended solely on the economic theory. In actual fact, this

[8] In this case the support space is symmetrically constructed but not centred on zero, since negative values are not admissible values for the economic theory.
[9] For the *Agriculture, Animal Husbandry and Fishing* the support was fixed at [0; 5.08]; for *Energy and Mining products* at [0;1.36]; for *Industrial products* at [0; 1.41]; for the *Construction* at [0;1.44]; for the *Market Services* at [0;2.88]; for the *Non Market Services* [0;6.18]. The general rule proposed could be to fix the highest bound of the efficiency parameter at a value equal to twice the calibrated parameter. This rule is considered valid unless the specified support includes the value 1 for the efficiency parameter. Recalling that a change in the parameter α changes the output in the same proportion for any given set of inputs, it is reasonable to hypothesize that the support for each branch includes values both lower and higher than one, to allow different neutral efficiency degree for each branch. For this reason we will include in all the supports the value 1 for the efficiency parameters

parameter represents the share of each input employed in the production process and therefore it must range between 0 and 1. The calibrated values, shown in Table 4.9, clearly confirm the theory and they will be used to define the prior probability distribution (see paragraph 4.4.2).
We have no prior knowledge either from the RSAM or other studies on the specific economic context of Sardinia concerning the *substitution parameter ρ*. Moreover most of the existing studies concerning CGE models, both at regional and national level, define the degree of substitution between these two inputs (i.e. value added and intermediate consumption) according to a Leontief technology, meaning that no level of substitution is allowed for them.
However, we can take advantage of the economic theory since it is well-known that the admissible values are all equal or greater than -1. In the choice of the supports, it has been preferred to define wide supports in order to have less impact on the estimates. To be precise, the sensitivity analysis is conducted considering four type of support spaces, which are greater than 50 times, 10 times, 5 times and 2.5 times of the minimal admissible value for ρ[10], and therefore it has been defined as [-50;50], [-10;10], [-5;5] and [-2.5;2.5] respectively. These supports are centred on zero and symmetrically constructed. The choice to include also negative values lower than -1, even though the economic theory states that σ should be greater than zero, has been done in agreement with the GME theory, that suggest to define a symmetric support centred on zero, so that if the data does not add information the estimated values converge to zero. It is worth to note that, for this parameter, also a value equal to zero takes an important meaning. Recalling the already known relation between the ρ and the substitution elasticity σ, a value of ρ equal to zero leads to a unitary elasticity of substitution and therefore the isoquants of the CES production function take the forms of those of the CD production function; this means that there is a unitary elasticity of substitution between the two inputs.
Our specification choice is also confirmed by the results presented by Mittelhammer and Cardell (1996) which have argued that for wider support spaces, GME results were quite similar to OLS results for moderate to large sample sizes and provided some degree of improvement over OLS for small samples.
The chosen support space for ρ might be different across the branches, also reflecting the different economic nature of each of them. It is worthwhile to note that in the choice of the final models, summarized in Table 4.10, for this estimation step, the adopted criteria consist of both considering the expected values, the admissible theoretical values and the observed sensitivity of the estimates to the support changes.
The results of this step of the estimation approach are shown in detail for each branch, in the Appendix 1. Along with the value of estimated parameters, each

[10] In a similar way, Golan et al (2001) decided to choose the support spaces of the unknown parameters in their AIDS demand system for Mexico. Based on a literature review on applied AIDS demand functions they found that the estimated coefficients on log prices were within the interval (-0.2;0.2) and that the intercepts and coefficients on log expenditure were within the interval of (-1;1).

model includes some diagnostic measures. More specifically, it has been determined the NE both for the entire model and for each of the estimated parameters.
The measure for each parameters can be used to evaluate the information in each of the variables, since it reflects, as argued by Golan (2008) the relative contribution (of explaining the dependent variable) of each one of the independent variables. In addition, according to Fraser (2000), values of the NE near to one mean that the solution is nearly uniform and that the data agree with the prior[11]. As shown in the tables 4.10, the obtained values of the NE for each model validate the support spaces choices.
To summarize, it is possible to state that for the *Energy and Mining products* the estimated value of the substitution parameter are not sensitive to the support specification. In these case the chosen supports was this with the lowest NE for this specific parameter. A quite similar logic was followed for the *Agriculture, Animal Husbandry and Fishing*. The estimated values of the substitution parameter were robust in magnitude to the change of the supports even if the estimated signs changed. In this case, the choices carried out agree with the support space which was at the same time the widest and with the lowest NE for the substitution parameter.
For the *Industrial product* and the *Market service* branches the estimated substitution parameters although they show the correct signs they express an estimated magnitude quite sensitive to the support specification. In these cases, the choice carried out lean towards those with both wider support space, lower associated NE and whose estimated results agree with the economic theory . This accounts for our lack of prior knowledge about the parameter values as mentioned before, as well as our efforts to let the data speak.
It is worth mentioning that for the *Construction* branch a more carefully choice was carried out. By following the already mentioned guidelines and the specified diagnostic tool the [-10;10] support would be chosen. However, with this support the related values of the elasticity of substitution would be a "non plausible" value. For this reason it was decided to choose the support [-5;5] whose diagnostic measure were almost identical with those of the [-10;10]. In this case it is important to emphasize the property of the GME estimation approach since, by taking advantage of its principles it makes possible to obtain the estimates also with one piece of information only. The sensitivity analysis for the *Non Market Services* branch was carried out until the support [-1;3], which is the chosen support, since the estimation process did not "take out" any information about the substitution parameter from the data.

[11] This diagnostic measure, as stressed by Golan et al (1996a) can be also used in the selection of the variables in a regression model. That is, a variable is extraneous in a regression model if its normalized entropy statistics is greater than 0.99. In this case the variables with a greater value of the NE are excluded from the regression model.

Table 4.10 – B-GME Estimates for the CES production function

	α	δ	ρ	σ	NE
Agriculture, animal husbandry and fishing[A]	2.546 (0.9999)	0.502 (0.9999)	-0.009 (0.99996)	1.009	0.9999
Energy and mining products[B]	0.736 (0.9951)	0.728 (0.8638)	-0.346 (0.9880)	1.531	0.9490
Industrial products[C]	1.354 (0.2857)	0.118 (0.5707)	-0.456 (0.9791)	1.840	0.6118
Construction[D]	1.039 (0.8721)	0.443 (0.9920)	-0.228 (0.9987)	1.295	0.9543
Market services[E]	1.837 (0.9519)	0.521 (0.9988)	-0.210 (0.9989)	1.266	0.9832
Non market services[F]	2.480 (0.9755)	0.611 (0.9688)	0.501 (0.9608)	0.666	0.9683

The normalized entropy value for each parameter is provided in the parentheses. Parameter supports:
A: α [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-10;10]; [-4.27;4.27] for the error
B: α [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-2.5;2.5]; [-0.56;0.56] for the error
C: [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-2.5;2.5]; [-2.59;2.59] for the error
D: α [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-5;5]; [-1;1] for the error
E: [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-5;5]; [-1.5;1.5] for the error
F: [0; 2times the value of the calibrated parameter]; δ [0;1] ρ[-1;3]; [-0.77;0.77] for the error
Source: our elaboration on the RSAM for Sardinia (2001)

Table 4.10 summarizes the estimated parameters as well as the obtained substitution elasticities between value added and intermediate consumption.
Some interesting comments arise from the results reported in Table 4.10. This *second estimation step*, characterized by adding a little more prior information on the unknown parameters, leads to estimates that are of the right signs (which is imposed for the efficiency and the share parameter) and whose magnitude (restricted between zero and one only for the distribution parameter δ) is in keeping with the economic theory. It is worth to note that the support for the α was constructed bearing in mind the Leontief fashion which is the "natural" underlying assumption of a RSAM. However, the results obtained were not considered the final results since different remarks could be underlined.

First of all, the estimated values for the elasticity of substitution confirm for all the sectors and explain well the existence of a CES relationship between the two factors of production.
Secondly, concerning the efficiency parameters, the estimated values highlight a good technology level for all the production branches (the "state of technology" in this case refers to the two input value added and intermediate consumption). However some estimates, namely for the *Industrial products*, *Construction* and *Market services* branches are relatively shift towards the upper specified bounds and relatively far away the "calibrated" values. This evidence could be the "clue" of a greater efficiency than those expected by considering only the RSAM structure. At the same time it might be necessary to consider a wider support space for the efficiency parameter in order to avoid that the estimated values of

the other parameters could be bias from a "wrong" specification of the α supports. This choice has been carried out also bearing in mind that a wrong specification of the supports might influence the values of the other estimated parameters. On this issue, Fraser (2000) highlights the need to be aware of the trade-off between imposing a priori information on the calculation process and simply letting the data speak. However, the same author underlines that this is not an easy task especially when estimating non-standard problems for which there is little information or knowledge to guide the estimation process.
The NE values for each parameter give also rise to some interesting considerations. For all the branches, the parameter that contributes more to explain the regional output (which is the dependent variable in this case) is the efficiency parameter α. On this issue, the best contribute is recorded for the *Industrial product*s branch, for which the value of the α is equal to 1.354 and the associated NE is equal to 0.2857.

The chosen supports for the substitution parameter will directly enter the following estimation step, namely the *third estimation step* represented by the GCE estimation.
This following *third estimation step* aims to extend and improve the results obtained in this second step. In particular, the GCE estimation process will consider a wider support for the efficiency parameter and a support for the substitution parameter which coincide with the support space which fit better with the data in this estimation step.
The third estimation step is required in order to test again and in a more general framework (which is not necessary strictly related to the RSAM structure) the relationship between intermediate consumptions and value added. This is done, considering the key role recovered by the substitution elasticity in the RCGE computation process and in the related economic simulation procedures.

4.4 The Third Estimation Step: a GCE Estimation

4.4.1 The optimization problem

Considering the extremely "limited" amount of information at our disposal and in order to improve the previous obtained estimates, the GCE estimator was used in order to formally introduce some "*a priori*" information on the unknown parameters in the optimization problem, both in terms of bounds of the economic admissible value for each of them, and in terms of their size of magnitude too. On this issue, regarding the definition of the support spaces and the prior probability distribution, we will take advantage both of the economic principles and the results obtained in the GME and B-GME estimation steps.
In particular, the previously obtained results, namely the results obtained in the *first estimation step* for the efficiency and the distribution parameters and the results obtained in the *second estimation step* for the substitution parameter, represent a good starting point for the evaluation of the magnitude of the unknown

parameters. Therefore, this information together with the results obtained in the previous estimation step where a sensitivity analysis was carried out enabled us to define a "more informative" support.
In this *third estimation step,* a more informative set of bounds will be used for each of the unknown parameters together with "a priori" information which will appear explicitly in the objective function in the form of a distribution probability.
Specifically concerning the support spaces for the three unknown parameters we proceeded in the following way.

Concerning the efficiency parameter α, the lowest bound was fixed at zero as required by the economic theory[12]. The highest bound was chosen by considering the obtained calibrated efficiency parameter and the results obtained at the end of the second stage of the estimation process.
Bearing this in mind, our choice was to fix the highest bound for all the branches equal to a value 10 times greater than the calibrated value. By doing so, in the support space we included the values obtained and chosen as the final results in the second step as well.
The choice to expand the support space from the observed values was carried out in alignment with other existing papers such as in Golan et al. (2001) and Rezek and Campbell (2007) who fixed support spaces which were wide enough to be non-binding.
A similar approach has been carried out by Howitt and Msangi (2006) who used values from a calibrated optimization model to ensure that the supports are centred on values that are feasible solution to the data constraints, and consistent with prior parameter values[13]. Moreover, by expanding the support space up to ten times the calibrated values we insert the estimation problem in a more general RCGE framework and which could mean to have different efficiency degree of each branch with respect to those calibrated from the RSAM.

Our choice concerning the definition of the support spaces of the *distribution parameter δ*, depended solely on the economic theory. In actual fact, this parameter represents the share of each input over the output and therefore it must range between 0 and 1. The calibrated values, shown in Table 4.9 will be used to define the prior probability distribution as well.
Concerning the substitution parameter support spaces, we introduced in the estimation process the support space which allows us to "extract" the greatest amount of information from the data, with refers to the second estimation step.

[12] In this case, the support space is not constructed symmetrically
[13] Howitt and Msangi (2006) stress that this specification of support values differentiates their approach with other GME production analysis used in the literature. In fact, the empirical GME literature says very little about how a set of feasible and consistent support values are defined for several interdependent parameters.

The chosen supports, which are different among the branches, are constructed considering the minimal admitted value for this parameter[14].

4.4.2 The prior probability distribution: a ME estimation

In this estimation step, it is our intention to use the calibrated information introduced in paragraph 4.3.2, in order to obtain a prior probability distribution on the "focused" defined support spaces for each branch.
However, there is a clearly ill-posed problem with the available data since we have only one piece of information (i.e. the calibrated parameter value) and we want to determine the related probability distribution which is the most likely set of relative frequencies which could have been generated in the greatest number of ways and is consistent with "what we know" (Golan et al, 1996a).
As this observed value was taken from the RSAM, we consider this information as a pure moment and specifically, due to the structure of the aggregate RSAM, as the first observed moment of the defined random variable whose realizations are the specified support points. When choosing the support space, first we considered the guidelines coming from the economic theory and the evidence resulting from the RSAM. Then, we decided to specify a support space, which only included plausible values for the studied parameters.
In particular, for the α parameter the different specified support spaces had a lowest bound equal to zero (the lowest bound was identical for all the branches) and a highest bound equal to a value ten times greater than the calibrated parameter. Once the support spaces were defined for each branch, we estimated the most likely probability distribution that could have generated the observed value with the ME estimator.
In order to find the M-dimensional (with M equal to 5) unknown probability distribution (which enables us to obtain the observed value as the expected value of the random variable), we used the ME estimator, firstly introduced by Jaynes (1957a; 1957b), according to which the entropy is maximized subject to the constraints represented by the data and the normalization requirement.
In the aim of explaining the applied procedure, a short description of the optimization problem will be carried out for the ME estimated probability distribution concerning the efficiency parameter of the *Agriculture, animal husbandry and fishing* branch.
From the RSAM, we calibrated the efficiency parameter and we obtained a value equal to 2.54, as shown in paragraph 4.3.2. This value represents the sole information used and the related optimization problem may lead us to modify the distribution of probabilities defined on a space of five points.
Therefore, we applied the ME estimation method in order to determine the probability distribution which generated the observed value. It is worth noting that by applying this estimator we will choose the probability distribution which is consistent with the data.

---

[14] Considering the relationship between substitution elasticity and ρ, our hypothesis is extremely valid since a value of the substitution parameter close to the highest bound leads us to a value of substitution elasticity close to zero, which corresponds to a Leontief substitution type.

The obtained probability distribution then entered the GCE estimation process as “the prior probability distribution”. Clearly, the “prior” expected values of the random values, defined for each unknown parameter, give back the calibrated value obtained from the RSAM, thanks to this probability distribution. Whenever this happens even after the estimation process has been completed it would mean that the chosen production technology and the chosen functional form describing the production system agree with the RSAM framework.
Proceeding in this way, the following step, namely the GCE estimation process, is aimed at minimizing the entropy distance between the data in the form of **p** and the estimated prior probability distribution in the form of **q**.
As stressed by Golan et al. (1996a) this is the underlying principle of the probabilistic distance or divergence. According to Good (1963), the optimization problem minimizes the cross-entropy between the probabilities that are consistent with the information in the data and the prior information **q.**

Below, Tables 4.11 and 4.12 showing the estimated probability distribution for α and δ are reported, respectively.

Table 4.11 - Estimated prior probability distribution for the efficiency parameter

	Prior probability distribution for α
Agriculture, animal husbandry and fishing	[0.710;0.208;0.061;0.018;0.005]
Energy and mining products	[0.712;0.206;0.060;0.017;0.005]
Industrial products	[0.713;0.207;0.058;0.018;0.005]
Construction	[0.710;0.207;0.060;0.018;0.005]
Market services	[0.711;0.207;0.060;0.017;0.005]
Non Market services	[0.710;0.207;0.060;0.018;0.005]

Source: our elaboration on the RSAM for Sardinia (2001)

As argued by Fraser (2000) “a priori” information alters the uniform distribution and replaces it with another distribution which describes the given parameter better.
In Table 4.11 we can see that the specified distribution for the efficiency parameters is an asymmetric distribution that places less probabilistic mass on the right tail. For the distribution parameters, whose support bounds are the same for all the branches, the specified distribution reflects the prior value of each specific branch.

Table 4.12 - Estimated prior probability distribution for the distribution parameter

	Prior probability distribution for δ
Agriculture, animal husbandry and fishing	[0.069;0.107;0.166;0.258;0.400]
Energy and mining products	[0.296;0.237;0.191;0.153;0.123]
Industrial products	[0.387;0.256;0.170;0.112;0.074]
Construction	[0.270;0.229;0.195;0.165;0.140]
Market services	[0.145;0.168;0.196;0.227;0.264]
Non Market services	[0.085;0.122;0.176;0.253;0.365]

Source: our elaboration on the RSAM for Sardinia (2001)

4.4.3 The estimated values and the elasticity of substitution

As already mentioned, this third estimation step was characterized by the introduction of "a priori" information both in terms of a "restrictive" set of bounds for the unknown parameters and in terms of ME estimated prior probability distribution for the efficiency and the distribution parameters. In this way, the prior information appears explicitly in the objective function.

Referring to the substitution parameters ρ it is important to note that although the supports introduced come from the previous estimation step, they combine with an underlying uniform probability distribution that places equal mass to each part of the distribution.

The results obtained, which can be considered the final results of our estimation problem, confirm the validity of our choices both concerning the three-step estimation procedure and the functional form selected to describe the production relationships.

In particular, the results of this three-step estimation process (shown in Table 4.13 below) improved the already "good" results. Figure 4.1 shows the isoquants which describe the estimated CES production function for each branch in question.

From an economic point of view, the parameter estimates comply with our expectations and appear to be consistent with the Sardinian economic context.

Table 4.13 - GCE Estimates for the CES production function

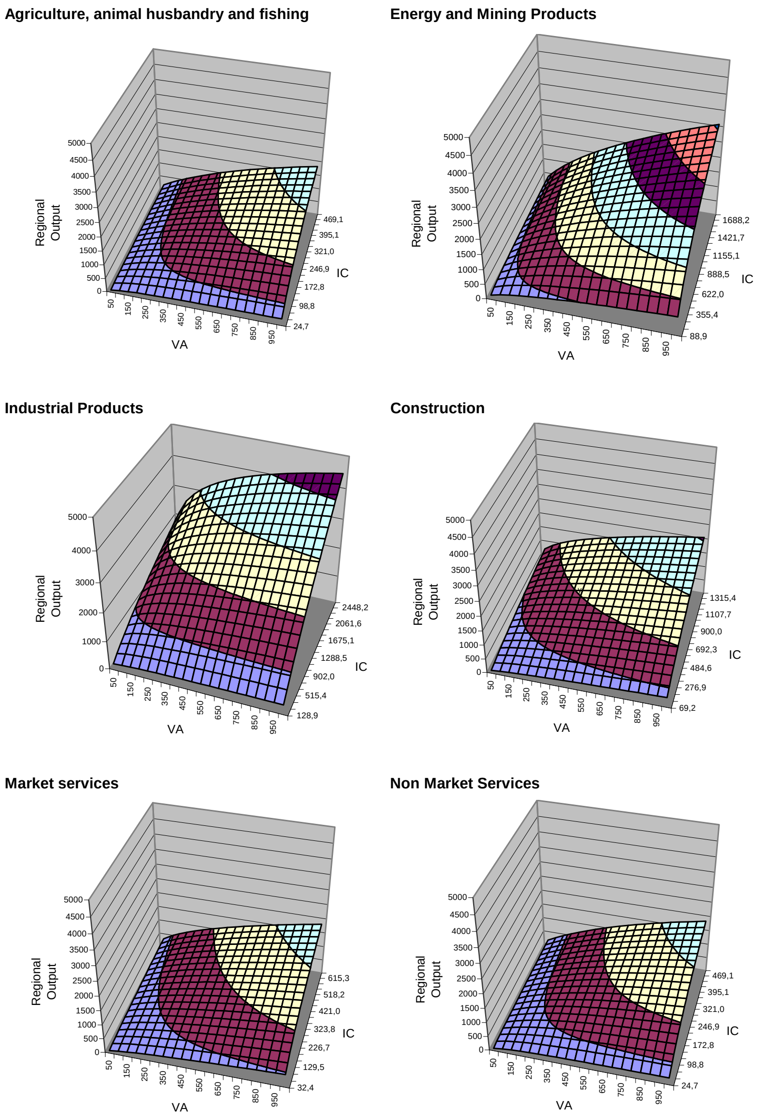
GCE Model	α	β	ρ	σ	NE
Agriculture, animal husbandry and fishing[A]	2.617 (0.5274)	0.704 (0.8923)	0.274 (0.9985)	0.785	0.806
Energy and mining products[B]	2.383 (0.9704)	0.552 (0.9930)	0.298 (0.9911)	0.770	0.9849
Industrial products[C]	2.068 (0.9904)	0.178 (0.7132)	0.189 (0.8015)	0.840	0.8350
Construction[D]	1.667 (0.8079)	0.397 (0.9733)	-0.089 (0.9998)	1.097	0.9270
Market Services[E]	2.149 (0.6507)	0.574 (0.9865)	0.046 (0.9999)	0.956	0.8790
Non Market Services[F]	2.287 (0.4324)	0.722 (0.8717)	0.758 (0.9909)	0.569	0.765

A: α [0;10 times the calibrated value]; δ [0;1] ρ[-10;10]; [-4.27;4.27] for the error component
B: α [0;10 times the calibrated value]; δ [0;1] ρ[-2.5;2.5]; [-0.56;0.56] for the error component
C: α [0;10 times the calibrated value]; δ [0;1] ρ[-2.5;2.5]; [-2.59;2.59] for the error component
D: α [0;10 times the calibrated value]; δ [0;1] ρ[-5;5]; [-1;1] for the error component
E: α [0;10 times the calibrated value]; δ [0;1] ρ[-5;5]; [-1.5;1.5] for the error component
F: α [0;10 times the calibrated value]; δ [0;1] ρ[-1;3]; [-0.77;0.77] for the error component
Source: our elaboration on the RSAM for Sardinia (2001)

Concerning the efficiency parameter, the *Agricultural, animal husbandry and fishing* branch appears to be the most efficient branch with an estimated coefficient equal to 2.617. However, efficiency parameters higher than one, as stressed by Ferrari and Manca (2008) seem to adequately reflect the technological structure of the Sardinian production system, characterized by small-medium enterprises, with a significant degree of technical flexibility.
Concerning the relative importance of the two aggregate inputs, namely value added and intermediate consumption, there is evidence for almost all sectors of the prevalence of value added over intermediate consumption due to the capital-intensive orientated type of activity of the island. Moreover, the results obtained confirm the calibrated measure introduced as "a priori" information concerning the distribution parameters.
Agriculture, Animal Husbandry and Fishing and *Non Market Services* represent the two branches with the highest value of distribution parameters. Thus, to say that the value added in these branches make up over seventy per cent of the total input; with the two estimated values equal to 0.704 and 0.722, respectively. On the other hand, the *Industrial product* branch represents the branch with the lowest value of the distribution parameter, that is to say in this branch there is the highest percentage of the intermediate consumption. For this reason, it is important to emphasize that this branch includes the Food product, beverage and tobacco industries, as stated in Chapter 3. These industries are usually characterized by a very high level of intermediate consumption (mostly coming from the agricultural activities). In fact, in this case the Food product, beverage and tobacco industries show an intermediate consumption coming from the

Agricultural, animal husbandry and fishing branch which represents over 35 per cent of the Sardinian total intermediate consumption.

Figure 4.1 – CES production functions for the six branches



Source: our elaboration on the RSAM for Sardinia (2001)

The obtained values for substitution elasticities concerning each branch illustrate and confirm the existence of a CES relationship between the two factors of production, value added and intermediate consumption[15]. This relationship is

[15] It is important to note that according to the seminal work of Arrow et al. (1961) the Sardinian CES function was implemented by considering aggregate quantities, namely value added and intermediate consumption, and therefore the elasticities endogenously obtained were aggregate

established for every sector, but it has a peculiarity which is important to note for the Construction branch whose isoquant appears quite similar to a Cobb-Douglas production function. For this branch, a 1 per cent change in the MRTS leads to a 1.097 per cent change in the input mix in order to keep the output at the same level.

For all the other branches, the substitution elasticities are lower than one. The lowest value is reached by the branch of *Non Market services* with a value equal to 0.569. This value being close to the lowest admissible value for the substitution elasticity highlights that for the Non Market services there is a low degree of substitution between value added and intermediate consumption, close to the Leontief functional form. This can be also seen by observing the actual share of the value added over the intermediate consumption. This value of the substitution elasticity makes it possible to state that a 1 per cent change in the MTRS leads to a 0.569 per cent change in the input mix, with the aim of maintaining the output constant. This is the confirmation of a quite high rigidity of the substitution degree between value added and intermediate consumption in the Non Market Service branch.

4.5 Some remarks on the proposed *three-step* estimation approach

The *three-step Entropy estimation approach* proposed in this study allows us to obtain estimates of the CES unknown parameter production functions for the region of Sardinia. The final obtained estimated parameters confirm the sign and magnitude required by the economic theory.

Figure 4.2 shows in detail the main aspects of the proposed *three-step* estimation approach.

However, several improvements could be introduced concerning both a "better" specification of the support spaces for all the unknown parameters and additional diagnostic tools.

At this stage of the research, it is our intention to give an important more general to this estimation approach, in the context of RCGE or CGE models. As stressed in Chapter 1, the determination of the behavioural parameter values, concerning both the production and the demand functions, represents one of the most important obstacles for a correct and coherent computation of a CGE model, since not all the unknown parameters can be calibrated from the RSAM.

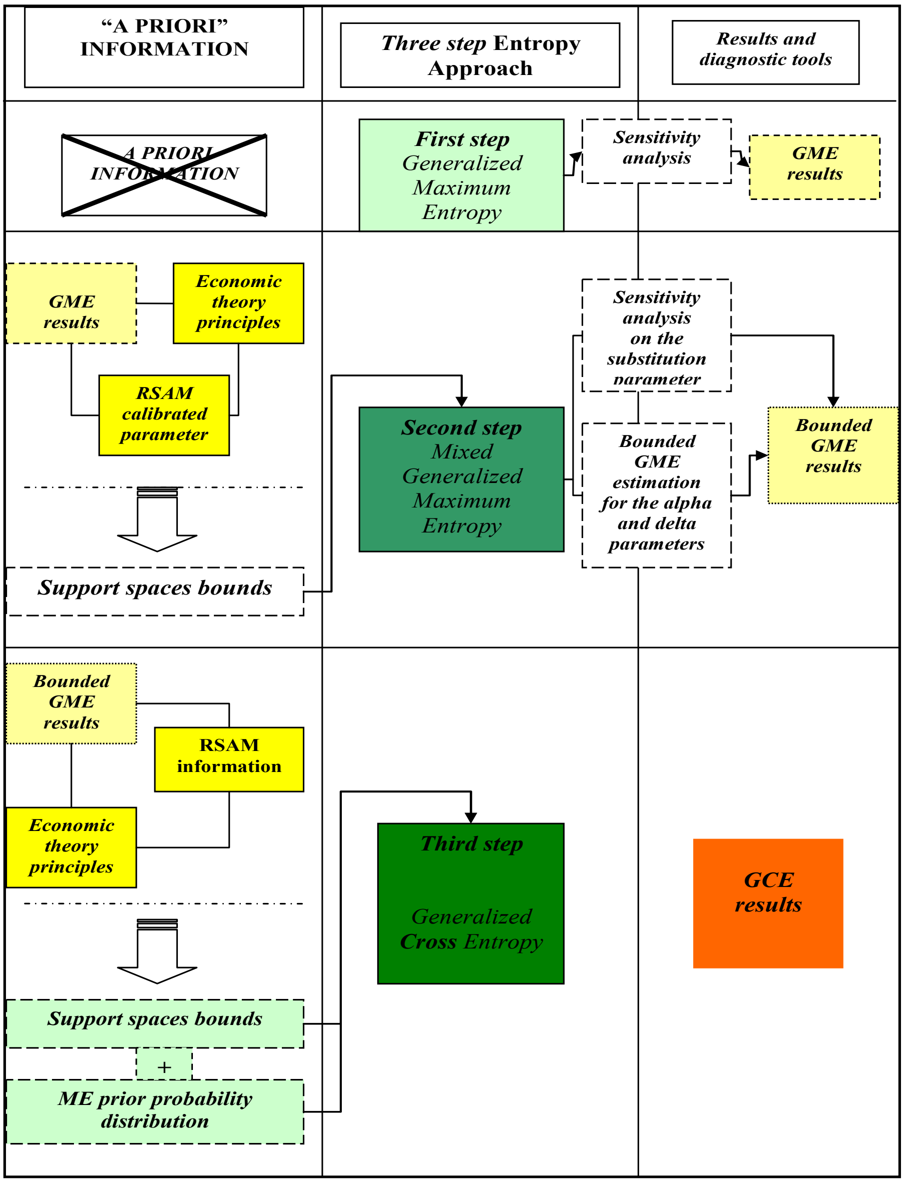
The proposed *three-step approach* might be used by RCGE or CGE modellers to merge two types of information which i) is taken directly from the RSAM/SAM and ii) which could be "extracted" from the RSAM/SAM. In agreement with the related economic theory, this could be done in order to reduce the complete uncertainty concerning the unknown parameters of the chosen functional form[16].

---

measures as well. Miyagiwa and Papageorgiou (2007) provide an interesting study concerning the endogenous aggregate elasticity of substitution.

[16] In this way, by taking advantage of the GME/GCE potentiality, it could be possible to obtain the value of the unknown parameters, by using the macro-information contained in the RSAM/SAM only.

Figure 4.2 - Three step entropy approach: main characteristics



To be clear, in order to completely specify the CES production function in a RCGE model framework, we need to know the values of the efficiency, distribution and substitution parameters. On one hand, the α and δ parameters can be calibrated from the RSAM, so that the model reproduces the benchmark dataset as an equilibrium solution. On the other hand, it is not possible to calibrate the value of the substitution parameter ρ from the same matrix. Therefore, we would need to obtain this parameter values through an imputation or an "external source" estimation procedure which has been the common practice up to now. Our idea is to obtain all the unknown parameter values through a simultaneous estimation process based on the RSAM information only. In actual fact, the

proposed estimation procedure enables us to introduce "what we know" from the RSAM (specifically concerning the efficiency and the distribution parameters) and at the same time to use this information to reduce the "complete uncertainty" which characterizes the substitution parameters.
The obtained results, although lacking in asymptotic statistics, confirm our efforts in translating an "ill-posed" estimation problem into a "well-behaved" problem.
It is worth noting that neither asymptotic nor finite sample inference on model parameters is provided here. As stated by Tonini and Jongeneel (2008), the main reason for this lack is the extremely small sample size (in some cases equal to one or two observations) which makes it unlikely that the sampling distribution of the estimator is similar to the asymptotic distribution. A possible solution, which could represent the natural development of this research, would be to use bootstrapping techniques to provide estimates of the finite sample distributions of parameters.
In view of this, when discussing the results we focused on evaluating the information content of the parameter estimates using the NE measure for both the signal and the whole system[17]. In particular, as already shown the NE measure helps us to assess the three different steps of our estimation strategy, by determining the change in information due to the change of support spaces.
Concerning the production side and particularly the important role played by the substitution elasticity in the RCGE context, the proposed approach achieves an important result since it allows the modellers to obtain the substitution elasticity by using the RSAM information and therefore in a "self-contained" approach. Traditionally, as stressed in Chapter 1, this solution has not been adopted up to now, due to data limitation and to econometric problems that usually lead to an "*ill-posed*" situation.

[17] As underlined by Lansink (1999) it is not uncommon to find high values of the NE.

Chapter 5

GME Approach for the Estimation of the Working-Leser Demand Functions

5.1 Specified Functional Form and Estimation Strategy

In order to analyze the household consumption trends, coming from the Sardinian RSAM, we estimated Engel curves using the functional form proposed by Working (1943) and Leser (1963), introduced in a general version by the [3.8] and [3.10].

It is important to note that the assumption of log-linearity in the relationship between budget shares and household expenditure, which underlies this functional form, agree with the specific studied context since a RSAM provides comprehensive information on variables regarding a particular year.

Concerning the consumption sphere, the RSAM describes the distribution of total household expenditure among the different branches in the specific period for which the matrix has been constructed. The benchmark year in the RSAM for Sardinia is the year 2001.

Therefore, it is absolutely plausible to pay no attention to any variations in the price level. For this reason, prices are considered constants and the only variables that could influence the consumption is the income, usually represented by the total expenditure.

The same evidence was noted by VanDriel, Nadall and Zeelenber (1997) and Beneito (2003) who stated that in this type of analysis it is unnecessary to introduce price variables in the equation system.

Generally, the main interest in the study of a demand system and furthermore in a demand system within a RCGE framework is represented by obtaining expenditure elasticity related to the different household groups in question (Balcombe et al., 2003; Browne, Ortman and Hendrick, 2007; Castaldo and Reilly, 2007).

In actual fact, consumer expenditure patterns and above all the estimates of expenditure elasticity can be a valid indicator of the "responsiveness" of the demand and therefore of the structure of the economy examined.

Bearing these considerations in mind, the main objective of this phase of our research was to estimate expenditure elasticity, concerning the demand of the

Sardinian households directed towards the branches included in the RSAM, which were suitable for the RCGE model equations.
Once again, we estimated expenditure elasticity in a "self-contained" approach, that is to say by using the RSAM information only. To our knowledge there are no other existing studies which have used the same approach.
For the estimation procedure, we took advantage of the GME estimation method in order to solve the problems caused by an "ill-posed" situation.
We estimated a demand equation system, for each of the six household income groups and for the branches of the Sardinia RSAM, without assuming any "a priori information".
The estimation process is based on a demand relationship equation system in agreement with the constraints imposed both by the consumer-theory and by the GME approach.
The aggregate RSAM matrix (see Table 3.4) proposed in this research illustrates branches, which are used here in order to analyze the distribution of the household consumption among different types of producer goods and services.
The available RSAM information allows us to differentiate estimates for different income groups and aggregate branches. As discussed in Chapter 3, the aggregate RSAM for Sardinia includes 6 income groups (or categories) of households.
It is useful to remember that households are classified on the basis of their annual income level, as follows:

i) Household income group A: disposable annual income lower than 9,300 Euro;
ii) Household income group B: disposable annual income from 9,300 to 12,400 Euro;
iii) Household income group C: disposable annual income from 12,400 to 15,500 Euro;
iv) Household income group D: disposable annual income from 15,500 to 24,800 Euro;
v) Household income group E: disposable annual income from 24,800 to 31,000 Euro;
vi) Household income group F: disposable annual income higher than 31,000 Euro;

The original structure of the RSAM where the production sphere is split into 23 branches (coming from the I/O table) was used as a basis for the estimation process, as already carried out for the applied production analysis. In actual fact, the branches shown in the original structure of the RSAM functioned as statistical units pertaining to each aggregate branch (see Table 3.1 for further details)[1].
Concerning the purposes of our analysis, the following branches were considered: *i)* Agriculture, Animal Husbandry and Fishing; *ii)* Energy and mining products; *iii)* Industrial products; *iv)* Market services; *v)* Non Market services.

[1] For example the branch of Agriculture, animal husbandry and Fishing is composed of the following two economic activities i) Agriculture, hunting and forestry and ii) Fishing and related services.

To be precise, our demand analysis is production branch orientated, according to the RSAM economic meaning and the RCGE framework.

A general specification of the Working-Leser equation system, for the problem in question, takes the form of a set of budget share equations, specified as:

$$w_{ji} = \alpha_{ji} + \beta_{ji} \ln(X_j) + \varepsilon_{ji} \quad [5.1]$$

where w_{ji} is the share of household *j* and branch *i* which is obtained as the ratio of the total expenditure on a particular branch *i* of the household group *j* to the total expenditure of the household group *j* analyzed. X_j indicates the total expenditure of the specific j^{th} household group studied.
The α_{ji} and β_{ji} are unknown parameters corresponding to the j^{th} household group and the i^{th} branch, while ε_{ji} is the error term that captures the unknown variation in the i^{th} share for the j^{th} household and for which standard econometric assumptions were made.
The index *j* will range from 1 to 6 while the index *i* from 1 to 5, reflecting the different household groups and branches, respectively.

Without loss of generality, it is possible to split the general expression which identifies the Working Leser demand system for the j^{th} household group. Obviously, the specified model, will be repeated for each of the six income groups.
The model for our particular purposes was composed of 5 equations, each one describing the i^{th} branch:

$$w_{j1} = \alpha_{j1} + \beta_{j1} \ln(X_j) + \varepsilon_{j1} \quad [5.2]$$

$$w_{j2} = \alpha_{j2} + \beta_{j2} \ln(X_j) + \varepsilon_{j2} \quad [5.3]$$

$$w_{j3} = \alpha_{j3} + \beta_{j3} \ln(X_j) + \varepsilon_{j3} \quad [5.4]$$

$$w_{j4} = \alpha_{j4} + \beta_{j4} \ln(X_j) + \varepsilon_{j4} \quad [5.5]$$

$$w_{j5} = \alpha_{j5} + \beta_{j5} \ln(X_j) + \varepsilon_{j5} \quad [5.6]$$

The model is also subject to the restriction that the sum of the shares must be equal to one, described as:

$$\sum_i w_{ji} = 1; \quad [5.7]$$

Moreover, the economic theory requires the following adding-up constraints, such that:

$$\sum_i \alpha_{ji} = 1; \quad [5.8]$$

$$\sum_i \beta_{ji} = 0; \qquad [5.9]$$

Clearly, the level of the expenditure elasticity in the Working-Leser functional form depends on the sign and magnitude of the estimated β coefficients, *ceteris paribus*. On this issue, the implemented model assumes that the weighted sum of elasticity for each group of households is equal to 1, as proved by Skorodova (1999) and Gradzewicz, Griffin and Zolkiewski (2006)

5.2 The GME Re-parameterization

In order to estimate the system of demand equations introduced by the equations from [5.2] to [5.5], we used the GME approach.
It is possible to state that the computational problem of this estimation process appears to be "complementary" to the one we had to deal with in the production context.
On one hand, it was necessary to deal with an economic problem where any "a priori" information on the parameter values was available from the RSAM. While it was not necessary to restrict the support bounds on particular admissible values, namely greater or lower than zero, since the economic theory states that both positive and negative values are possible for the two unknown Working-Leser parameters.
Concerning the magnitude of the parameters to be estimated, there are not to our knowledge, other existing studies which tried to estimate a complete demand system establishing the estimation process on the RSAM data only and within a RCGE context.
On the other hand, the consumer theory requires additional specific constraints such as the adding-up requirements over the unknown parameters, which must be verified for the entire demand system.
Having constructed the equation system and introduced the consumer-theory adding up constraints, the GME re-parameterization must be carried out.
It is worth noting that the estimation process was carried out in an extremely limited information approach. In spite of this, the GME approach properties allow us to obtain statistical estimates of the expenditure elasticity which agree with the approach proposed in this research and the economic context in question.
As stressed in Chapter 2, the GME objective maximizes the joint entropy of the parameters and the error terms. In order to write these two combined entropy measures which represent our "all-in-one" objective function it is necessary to express all the unknown coefficients and the error terms described in the general expression [5.1] in terms of proper probabilities which are defined over some specified support spaces.
For example, to transform the coefficient $\alpha_{ji,}$ we started by choosing a support space which is a set of discrete points $\mathbf{a}_{ji} = \left(a_{ji1}, a_{ji2}, \ldots, a_{jiS}\right)'$ for j=1,2,…J, and i=1,2,…I, where S≥2 identifies the number of support points which are uniformly

spaced out and which span the possible range of the unknown coefficients represented by the interval $[a_{ji1}, a_{jiS}]$.

Since we have no knowledge regarding the coefficients from the economic theory, we specified the supports to be symmetric around zero with "large" negative and positive bounds. This choice was made in agreement with the suggestion brought forward by Golan et al. (1996a), Golan (2008) and many existing studies such as Fraser (2000), Shen and Perloff (2001) and Nganou (2004).

The number of support points of the support spaces for each α_{ji} was fixed at S=5. Subsequently a vector of unknown weights $\mathbf{p}_{ji} = \left(p_{ji1}, p_{ji2,} ..., p_{jiS}\right)'$ where S=5 and such that $\sum_{s=1}^{S} p_{jis} = 1$ was introduced. By proceeding in this way, each unknown α_{ji} was re-written as $\sum_{s} a_{jis} p_{jis} = \alpha_{ji}$ for all *j* and *i*.

Similarly, the unknown parameter β_{ji} was re-parameterized using the same approach and therefore defining a set of discrete points $\mathbf{b_{ji}} = \left(b_{ji1}, b_{ji2}, ..., b_{jiM}\right)'$ and a vector of unknown weights $\mathbf{c_{ji}} = \left(c_{ji1}, c_{ji2,} ..., c_{jiM}\right)'$ for each β_{ji}, with M equal to 5 representing the number of points defined on the support space. In this way, $\beta_{ji} = \sum_{m} c_{jim} b_{jim}$ for all *j* and *i*.

The error term ε_{ji} was treated as unknown and re-parameterized following a similar procedure to the one introduced for the unknown parameters, so that $\varepsilon_{ji} = \sum_{z=1}^{Z} v_{jiz} r_{jiz}$. The transformation is done by specifying a vector of Z≥2 discrete points $\mathbf{v_{ji}} = \left(v_{ji1}, v_{ji2}, ..., v_{jiZ}\right)'$ distributed uniformly around zero and an associated vector of proper unknown weights $\mathbf{r_{ji}} = \left(r_{ji1}, r_{ji2}, ..., r_{jiZ}\right)'$ where $\sum_{z=1}^{Z} r_{jiz} = 1$.

With GME as stressed by Golan et al. (2001) there is no need to assume any subjective information on the distribution of the probabilities .

According to the GME theory, each support space and the associated probability distribution can be of different dimension.

In this study we used the same specification for all the α and the β parameters and more specifically for all the α and the β parameters pertaining to each branch and for all the household income groups.

5.3 The Optimization Problem and the Choice of the Support Spaces

The re-parameterized coefficients and error terms allow us to re-write (for the j[th] household group) the Working-Leser function described by the [5.1] as:

$$w_{ji} = \sum_s a_{jis} p_{jis} + \sum_m b_{jim} c_{jim} \ln(X_j) + \sum_z v_{jiz} r_{jiz} \quad \text{for i=1,2,...} \qquad [5.10]$$

The objective function of the related optimization problem is given by:

$$H(\mathbf{p},\mathbf{b},\mathbf{w}) = -\mathbf{p}'\ln(\mathbf{p}) - \mathbf{b}'\ln(\mathbf{b}) - \mathbf{r}'\ln(\mathbf{r}) =$$
$$= -\sum_{i=1}^{I}\sum_{s=1}^{S} p_{jis}\ln\left(p_{jis}\right) - \sum_{i=1}^{I}\sum_{m=1}^{M} b_{jim}\ln\left(b_{jim}\right) - \sum_{i=1}^{I}\sum_{z=1}^{Z} r_{jiz}\ln\left(r_{jiz}\right) \qquad [5.11]$$

subject to the data constraint represented (for each i) by the general expression [5.10] and the GME adding up conditions, such that:

$$\sum_s p_{jis} = 1 \quad \text{for i=1,…,5;} \qquad [5.12]$$

$$\sum_m b_{jim} = 1 \quad \text{for i=1,…,5 ;} \qquad [5.13]$$

$$\sum_z r_{jiz} = 1 \quad \text{for i=1,…,5;} \qquad [5.14]$$

Precisely, the GME estimator maximizes the [5.11] subject to the data constraints, represented by the general equation [5.10], the adding up GME requirements as expressed by [5.12], [5.13] and [5.14], the restriction for the shares to add to one and the consumer-theory restrictions described by [5.7], [5.8] and [5.9] respectively.

The solution to this maximization problem is unique. Forming the Lagrangean and solving for the first order conditions yields the optimal solution from which the following point estimates for our demand system are obtained, for each *i*:

$$\hat{\alpha}_{ji} = \sum_{s=1}^{S} a_{jis}\widehat{p}_{jis}; \quad \text{for i=1,…,5} \qquad [5.15]$$

$$\hat{\beta}_{ji} = \sum_{m=1}^{M} c_{jim}\hat{b}_{jim}; \quad \text{for i=1,…,5} \qquad [5.16]$$

$$\hat{\varepsilon}_{ji} = \sum_{z=1}^{Z} v_{jiz}\hat{r}_{jiz}; \quad \text{for i=1,…,5} \qquad [5.17]$$

As already mentioned, the GME estimator uses all the data points and does not require restrictive moment or distributional error assumptions. Therefore, the GME is robust for a general class of error distributions unlike the ML estimator, as stressed by Golan et al. (2001). In fact, the GME estimates as stated by these authors may be used when the sample is small, as in our estimation situation.
Without assuming any "external" information on the parameters, the intuition behind the GME approach, as stressed by Shen and Perloff (2001) is that the frequency that maximizes entropy is a reasonable estimate of the true distribution when we lack other information.

We will obtain the expenditure elasticities from the estimated parameters as described by [5.16], through the following expression:

$$\eta_{ji} = \left(\hat{\beta}_{ji} + w_{ji}\right)\frac{1}{w_{ji}} = \frac{\hat{\beta}_{ji}}{w_{ji}} + 1 = \frac{\sum_{m=1}^{M} c_{jim}\hat{b}_{jim}}{w_{ji}} + 1; \text{ for i=1,…,5} \qquad [5.18]$$

Since the values of the estimated α and β parameters have not a direct economic meaning, we will give more emphasis to the elasticity that are obtained from the parameters in our explanation of the results.

Concerning the support spaces, since no "a priori" information on the parameter values was available, the support vectors for both the α and β parameters must be wide enough to include all the possible outcomes, as already proved by Nganou (2004). Moreover, by doing so the impact of the support space on the parameters is reduced while that of the data is increased, as stressed by Golan et al. (1996a).
The specific support chosen for each household income group is the result of a sensitivity analysis that we performed by making a moderately large change in the support spaces while maintaining the same center of the support.
We also carried out this procedure in order to verify the sensitivity of the estimated coefficients on the support changes.
The change of the support space has a negligible effect on the estimated coefficients, which appeared to be extremely robust both in terms of sign and magnitude.
This stability also had a stabilizing effect on the expenditure elasticities, which were obtained from the expression [5.18].
In order to avoid repetition and as an example Table 5.1 shows the change in the value of the expenditure elasticity to the change of the support space[2] concerning the five different branches for the household group with the highest level of annual income.
When choosing the final models from which we obtained the expenditure elasticity values for the entire household system, we tended towards those with

[2] The estimated parameters appear to be extremely robust in terms of signs; the magnitude is extremely stable from the run 1 to the run 4 and relatively changes in the run 5 when a different support space between the α and the β are specified.

the lowest associated normalized entropy ratio, since by observing the results of the sensitivity test we observed that the estimates are highly robust to the support change. Therefore, among all the supports, we finally chose the one that "extracted" most of the information and suited the data better.

Table 5.1 - Expenditure elasticity for Household group F

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.5268	1.5268	1.5267	1.5240	1.1970
Energy and Mining products	2.7768	2.7768	2.7765	2.7680	1.6645
Industrial products	0.9570	0.9570	0.9570	0.9571	0.9839
Market services	1.0055	1.0055	1.0055	1.0055	1.0020
Non Market services	0.8963	0.8964	0.8964	0.8972	0.9613
Normalized Entropy Ratio	*0.99999*	*0.99999*	*0.99987*	*0.98743*	0.99986

Parameter supports:
GME 1 [-100;100] for α and β; [-1;1]for the error terms;
GME 2 [-50;50] for α and β; [-1;1]for the error terms;
GME 3 [-10;10] for α and β; [-1;1]for the error terms;
GME 4 [-1;1] for α and β; [-1;1]for the error terms
GME 5 [-10;10] for the α and [-1;1]; [-1;1]for the error terms
Source: our elaboration on the RSAM for Sardinia (2001)

The support space that added most information to the estimation process is represented by the support [-1;1] for the α and β parameters. These values appeared to be highly plausible in this specific case, since we dealt with the logarithmic transformation of the expenditure and the share of the total expenditure.
Therefore, for all the branches the chosen support was [-1;1] for both the α and β parameters,[3] while the support for the error terms was naturally bounded between -1 and 1 because all the dependent variables were relative shares lying between 0 and 1.
The relatively large support spaces specified for the re-parameterized parameters implied the absence of strong prior information from the supports on the parameters. The parameter estimates and their change to the change of support spaces are reported in Appendix 2, while the more interesting values of the expenditure elasticity will be discussed in the next paragraph.

[3] Golan et al. (2001) stressed that in a variety of AIDS which represent the natural "extension" of the Working Leser models both the estimated intercept and the estimated coefficients on log expenditure were within the interval of (-1;1).

5.4 GME Estimation Results: the Obtained Expenditure Elasticity

The Working-Leser model does not give a direct interpretation of the expenditure elasticity, since it is specified in the log-linear form. Therefore, the relevant elasticity cannot be directly accessed in the parameter estimates. For this reason, it is necessary to introduce the expression [5.18].

De Mello, Pack and Sinclair (2002) underlined the same evidence for the AIDS, which represents the natural extension of the Working Leser model.

From the results obtained in the previous steps, we obtained the final expenditure elasticity. These measures, as stressed by several authors such as Akbay Boz and Chern (2007) are used to approximate corresponding unconditional income elasticities.

Table 5.2 shows our estimated expenditure elasticity, which refer to the five aggregate branches.

It is important to emphasize that the computed expenditure elasticity reflect the elasticity of demand towards "production branches" and are not directed towards "consumer goods" as they usually are. This evidence generated deep reflection concerning their economic meaning and their interpretation.

Table 5.2 - Expenditure elasticity for the six household groups of the Sardinian RSAM

	Household A	Household B	Household C	Household D	Household E	Household F
Agriculture, animal husbandry and fishing	1.3289	1.3569	1.4412	1.4133	1.5859	1.5240
Energy and Mining products	3.5535	2.6142	2.9765	3.0417	2.4443	2.7680
Industrial products	0.9614	0.9627	0.9592	0.9633	0.9523	0.9571
Market services	1.0035	1.0024	1.0047	1.0044	1.0066	1.0055
Non Market services	0.9280	0.9590	0.9166	0.9195	0.9075	0.8972

GME parameter supports: [-1;1] for α and β; [-1;1]for the error terms:
Households are divided as follows: Household income group A: disposable annual income lower than 9,300 Euro; Household income group B: disposable annual income from 9300 to 12400 Euro; iii)Household income group C: disposable annual income from 12,400 to 15,500 Euro; Household income group D: disposable annual income from 15,500 to 24,800 Euro; Household income group E: disposable annual income from 24,800 to 31,000 Euro; Household income group F: disposable annual income higher than 31,000 Euro;
Source: our elaboration on the RSAM for Sardinia (2001)

As shown in Table 5.2 the models provide information on the tendency of households at different income levels to adjust their demand over the branches in relation to their income.

Expenditure elasticity for the *Agriculture, Animal Husbandry and Fishing* branch are quite high suggesting that the demand for this branch is expenditure elastic to the change in consumer expenditure. Moreover, the responsiveness to changes in

consumer expenditure grows with the level of income thus suggesting the incidence of some “luxury” products inside the branch, which would appear with the rise in income.
The products of the *Energy and Mining branch* seem to be expenditure elastic to the change of the total expenditure too. For example, concerning the household group C, a 1 per cent change in the total expenditure generates a 2.9765 per cent change in the expenditure towards this branch.
A lower degree of expenditure elasticity, close to one for all the branches, can be seen for the *Market services* which include economic activities such as Transport and communication, Banking and financial services.
The values of the expenditure elasticity suggest that Market services represent “necessary” services for all the households. Within this branch, the lowest expenditure elasticity is recorded for the household group with an annual disposable income level between 24,800 and 31,000 Euro. For this specific group a 1 per cent increase in total expenditure would increase the demand for Industrial products by 0.9523 per cent.
The tendency of the *Non Market service* elasticity is much more interesting, because the elasticity in all the studied income levels, which are less than one, tends to decrease with the rise in income. This suggests that for households with a higher income level, the percent change in the quantity of Non Market services requested is less than the percent change in the consumer expenditure, probably because they “substitute” some of these services with other services in the Market service branch.
It is important to note that the estimated magnitude of the expenditure elasticity for Industrial products agrees with our expectations due to the fact that the Food, tobacco and beverage industries are included in this branch and they have an important incidence on the expenditure share of the Industrial product branch. The value of expenditure elasticity lower than unity proves that these products are “necessary” goods for all the household income groups in question.
For example, a 1 per cent increase in the total Industrial Product expenditure for Household group F would increase the demand towards the Industrial product branch by 0.9571 per cent.
It is worth noting that all these estimates, imply that as the household’s total expenditure increases (resulting from an income increase), the demand towards the analyzed branches will also increase proportionally, more than proportionally or less than proportionally.

The obtained results represent a first attempt to estimate the expenditure elasticity in a RSAM context and within a RCGE model framework without considering any external information.
The estimates were obtained by using the GME estimation tool. Clearly, a different aggregation of the branches more focused on different types of “consumer goods” could be more attractive in a classical demand analysis framework. However, our research focuses on the aim of estimating behavioural parameters in a situation characterized by lack of data, which is a regional context.

The chief contribution of this research would be to show that it is also possible to estimate expenditure elasticities from a set of budget shares in a context where price information is assumed to be constant and where, statistically speaking, the lack of data usually leads to an “ill-posed” situation.

Concluding Remarks

The main objective of this research was to develop an estimation strategy which allows us to obtain production and demand function parameter estimates in a RCGE context where the available dataset is represented by the information contained in a RSAM only.

Without loss of generality, this estimation problem can be described as an "ill-posed" situation, since a RSAM does not provide sufficient information to enable us to achieve a consistent estimation process and consequently valid and significant estimates. In actual fact, working with a RSAM, or with a SAM at a national level, means to deal with a situation characterized by insufficient (even negative) degrees of freedom to be analyzed by using a conventional econometric approach.

Ill-posed problems are likely to occur in applied economics and in particular in the case of RCGE and CGE models where the benchmark dataset is represented by the SAM at regional or national level, respectively.

The role and the importance of CGE models are generally recognized, since they represent an efficient tool which is able to combine the theoretical general equilibrium with economic data, in order to provide statistical and economic institutions and governments with useful simulation instruments for policy evaluation.

Despite the widespread use of RCGE and CGE models, - due to the fact that they have been considered as an ideal bridge between economic theory and applied policy research (Bergman and Henrekson, 2003) – they have often been doubted for the weak econometric foundations upon which they are typically based (Jorgenson 1984; Shoven and Whalley 1992; McKitrick 1998). In particular, they are criticized for the selection of the parameters describing the behaviour of economic agents and consequently their validity, with particular regard to those related to production and demand spheres.

The implementation of a RCGE or CGE model requires different and interdependent phases whose results join together to achieve the main aims for which the model is constructed, usually represented by policy simulations, as above mentioned.

The phase that has frequently been questioned concerns those parameters whose values cannot be "calibrated" from the SAM, such as elasticity in both production and demand contexts, but for which additional external information is required.

Weak points found in literature concerning the two main approaches (namely imputation procedure and econometric approach) suggested to overcome the difficulties related to this phase, were the main motivation of our research and stimulated us to introduce a "self-contained" RSAM-based estimation approach,

in order to obtain the behavioural parameter values characterizing the production and demand spheres of the Sardinian RCGE model.
Estimates were carried out by using the RSAM for the Italian region Sardinia for the year 2001.
Firstly, starting from the 23-branch RSAM version we obtained an aggregate consistent version of the RSAM production system detailed into 6 branches.
Secondly, we chose the functional forms describing producer and household behaviours by studying the parameterizations which better fit the available data and agreed with the analyzed framework.
For the production sphere, according to the RCGE literature, a two-stage production process was modelled by using the CES functional form for both the specified levels, where the top-level refers to the value added and intermediate consumption which combine to obtain regional output, while the second level refers to capital and labour incomes which join together to obtain value added. The selected model gave more flexibility to the specified relationships since the CES function includes the Leontief (in an imputation context generally used to model the top-level relationships) and the CD specifications (in an imputation context generally used to model the second level relationships) as special cases.
In order to analyze the demand of Sardinian households for consumption we specified a system of equations based on the Working Leser functional form, according to the economic meaning of the RSAM which provides comprehensive information on variables regarding a particular year and therefore does not consider any variations in the price-level. By using this specification, we related the share of expenditure for consumption (pertaining to a certain branch) to the logarithm of the total household expenditure.

The proposed estimation approach, which was differently structured for production and demand functions, enabled us to obtain the values of the parameters for the CES production functions and the values of the elasticity related to the Working-Leser system – describing the demand responsiveness of Sardinian households regarding consumption – in a completely RSAM-based approach. Since no external information or already existing parameter values were used, the obtained values can be incorporated into the computation process of the RCGE model for Sardinia without any risk of biasing the results of policy simulations.
For the production side, the estimation strategy was constructed by using the GME and GCE estimators, since for all the analyzed branches we had negative or very limited degrees of freedom.
The most interesting aspect of our estimation procedure was that it enabled us to merge the phases of calibration and estimation. In actual fact, we simultaneously obtained estimates of the efficiency, distribution and substitution parameters of the CES function by using a procedure based on three steps, each of them characterized by a different degree of prior information.
In the first step, we did not assume any “a priori” information concerning the unknown parameters of the CES function. The obtained estimates confirmed the

existence of a CES relationship between the value added and the intermediate consumption for the branches in question.
However, we introduced a second estimation step in order to improve the estimates and by taking into account the information calibrated from the RSAM and from the production theory. This step, which we called a "B-GME", introduced the values for the efficiency and the distribution parameters obtained as results of the calibration phase into the estimation process. In agreement with the GME framework, we specified this information in terms of support bounds. The interesting estimated results were further improved by adding an "a priori" probability distribution to the parameter support (obtained by using the classical ME estimator and whose expected value was the calibrated parameter). This final step, characterized by the GCE method, enabled us to find estimates for all parameters which completely agree with the economic theory requirements and the Sardinian context.
Concerning the estimated values of the substitution parameters, from which we obtained the values of the elasticity of substitution - which specify the degree of substitutability between the two inputs, value added and intermediate consumption, considered in the production process – it is worth noting that they gave us the opportunity of highlighting some peculiarities of the Sardinian economic context. In particular, all the branches, except for the Construction branch, have substitution elasticity values lower than one. The branch of Construction has an elasticity of substitution close to one, which characterizes the CD production function. The branch which reaches the lowest value of elasticity and therefore a greater rigidity in value added/intermediate consumption substitution is the branch of the Non Market Services.
Referring to the elasticity of substitution between capital and labour incomes it is worth emphasizing that we obtained substitution elasticity lower than unity, which confirm that all the branches agree with the specified CES relationship. This relationship has a specificity for Energy and Mining product and Non Market service branches, whose elasticity values highlight a greater rigidity in capital/labour substitution than the other studied branches.
Concerning demand, an interesting aspect of the available RSAM was the subdivision of the households into six groups based on their annual disposable income level. This subdivision enabled us to estimate different Working Leser demand systems and therefore to obtain different interesting information for each household group.
An innovative issue of the demand system implemented was the analysis of the expenditure of the households for production branches. The results obtained in term of expenditure elasticity illustrate the different degrees of responsiveness of the household demand for consumption, on the basis of their annual disposable income level.
In particular, it should be noted that the products coming from the Agriculture, animal husbandry and fishing branch together with those produced by the Energy and mining branch recorded high expenditure elasticity values underlining that the demand is expenditure elastic.

On the other hand, products from the Industry branch and services pertaining to the Non Market service branch appear to be “necessary goods”; moreover, the responsiveness of the demand for these products decreases with the increase in income.

The obtained results gratify our efforts in transforming an “ill-posed” estimation problem into a “well-behaved” problem. Moreover, the proposed estimation strategies allowed us to obtain estimated parameter values which agree with the economic context in question and above all reflect the level of aggregation of the benchmark dataset.
To sum up, the proposed estimation procedures demonstrate that it is possible to completely parameterize the production and demand spheres of a RCGE model without using any external information.
Specifically concerning the production context, the obtained estimates are the results of a three-step procedure which is based on the inter-connection between the phases of calibration and estimation.

However, concerning the proposed strategy several plausible improvements can be made.
Firstly, some further diagnostic tools and inference procedures to assess the validity of the obtained estimates could be carried out. For example, bootstrapping techniques could be use to obtain confidence intervals for the parameter estimates.
Secondly, an interesting point to be developed and improved may concern the estimation process for the production side. In this study we performed estimates for the Sardinian production system separately for the top and the second level relationships. It might be interesting to carry out the estimation processes simultaneously by introducing the GME Instrumental Variable (IV) method, thus increasing the powerful of the estimates.
Furthermore, the performance of the three-step estimation approach might be analyzed for the demand side as well, if “ex-ante” consistent information was available or could be “extracted” from the RSAM.
In addition, a further development on this issue may concern a more comprehensive definition of the parameter support spaces which could be combined with the sensitivity analysis carried out by following the three-step estimation procedure proposed.
Finally, it would be interesting to extend the analysis to the sphere of commerce and to the other transactions of the economic agents.

Appendix 1

Table A1.1 - Sensitivity Analysis of B-GME estimates for Agriculture, Animal Husbandry and Fishing

	B-GME 1	B-GME 2	B-GME 3	B-GME 4
α	2.546 (0.9999)	2.546 (0.9999)	2.546 (0.9999)	2.549 (0.9999)
δ	0.501 (0.9999)	0.502 (0.9999)	0.502 (0.9999)	0.502 (0.9999)
ρ	-0.009 (1.000)	-0.009 (0.99996)	0.009 (0.9999)	0.008 (0.9999)
Entropy value	7.024	7.024	7.024	7.024
NE	0.9999	0.9999	0.9999	0.9999

*The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME4;
δ[0;1] from B-GME1 to B-GME4;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; for the error term: [-4.27;4.27];
Source: our elaboration on the RSAM for Sardinia (2001)

Table A1.2 - Sensitivity Analysis of B-GME estimates for Energy and Mining products

	B-GME 1	B-GME 2	B-GME 3	B-GME4
α	0.739 (0.9945)	0.739 (0.9946)	0.738 (0.9947)	0.736 (0.9951)
δ	0.728 (0.8648)	0.728 (0.8648)	0.728 (0.8646)	0.728 (0.8638)
ρ	-0.347 (0.9999)	-0.346 (0.9992)	-0.347 (0.9970)	-0.346 (0.9880)
Entropy value	5.683	5.681	5.678	5.664
NE	0.9531	0.9528	0.9521	0.9490

*The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME4;
δ[0;1] from B-GME1 to B-GME4;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; for the error term: [-0.56;0.56];
Source: our elaboration on the RSAM for Sardinia (2001)

Table A1.3 - Sensitivity Analysis of B-GME estimates for Industrial Products

	B-GME 1	B-GME 2	B-GME3	B GME 4
α	1.318 (0.6400)	1.356 (0.4506)	1.359 (0.2703)	1.354 (0.2857)
δ	0.420 (0.9840)	0.252 (0.8389)	0.162 (0.6794)	0.118 (0.5707)
ρ	-15.568 (0.9384)	-4.173 (0.8873)	-1.620 (0.9332)	-0.456 (0.9791)
Entropy value	11.669	11.160	10.945	10.851
NE	0.7733	0.6687	0.6276	0.6118

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME4;
δ[0;1] from B-GME1 to B-GME4;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; for the error term: [-2.59;2.59];
Source: our elaboration on the RSAM for Sardinia (2001)

Table A1.4 - Sensitivity Analysis of B-GME estimates for Construction

	B-GME 1	B-GME 2	B-GME 3	B-GME 4
α	1.014 (0.8921)	1.036 (0.8744)	1.039 (0.8721)	1.039 (0.8716)
δ	0.470 (0.9978)	0.445 (0.9925)	0.443 (0.9920)	0.443 (0.9919)
ρ	-8.840 (0.9804)	-0.878 (0.9952)	-0.228 (0.9987)	-0.0575 (0.9997)
Entropy value	5.346	5.273	5.267	5.266
NE	0.9567	0.9541	0.9543	0.9544

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME4;
δ[0;1] from B-GME1 to B-GME4;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; for the error term: [-1;1];
Source: our elaboration on the RSAM for Sardinia (2001)

Table A1.5 - Sensitivity Analysis of B-GME estimates for Market Services

	B-GME 1	B-GME 2	B-GME 3	B-GME 4
α	1.766 (0.9678)	1.827 (0.9543)	1.837 (0.9519)	1.844 (0.9502)
δ	0.504 (0.9999)	0.519 (0.9991)	0.521 (0.9988)	0.523 (0.9986)
ρ	-5.231 (0.9932)	-0.761 (0.9964)	-0.210 (0.9989)	0.041 (0.9998)
Entropy value	10.190	10.159	10.154	10.150
NE	0.9996	0.9833	0.9832	0.9830

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME4;
δ[0;1] from B-GME1 to B-GME4;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; for the error term: [-1.5;1.5];
Source: our elaboration on the RSAM for Sardinia (2001)

Table A1.6 Sensitivity Analysis of B-GME estimates for Non Market Services

	B-GME 1	B-GME 2	B-GME 3	B-GME 4	B-GME 5
α	3.146 (0.9998)	3.146 (0.9998)	3.146 (0.9998)	3.146 (0.9998)	2.480 (0.9755)
δ	0.572 (0.9869)	0.572 (0.9869)	0.572 (0.9869)	0.572 (0.9869)	0.611 (0.9688)
ρ	0.0002 (1.000)	0.0002 (1.000)	0.0002 (1.000)	0.0002 (1.000)	0.501 (0.9608)
Entropy value	9.156	9.156	9.156	9.156	8.922
Normalized Entropy Ratio	0.9956	0.9956	0.9956	0.9956	0.9683

The normalized entropy value for each parameter is provided in the parentheses.
Parameter supports:
α:[0; 2times the value of the calibrated parameter] from B-.GME1 to B-GME5;
δ[0;1] from B-GME1 to B-GME5;
ρ: GME1 [-50;50]; GME2: [-10;10]; GME3: [-5;5] GME4: [-2.5;2.5]; GME5: [-1;3] for the error term: [-0.77;0.77];
Source: our elaboration on the RSAM for Sardinia (2001)

Sardinia production system: production function described by the expression [3.7]

Results of the three-step entropy approach:

Table A1.7 - GCE estimates for CES production function (third estimation step)

GCE Model	α	β	ρ	σ
Agriculture, animal husbandry and fishing[A]	1.682 (0.5640)	0.490 (0.9997)	0.0095 (0.9990)	0.990
Energy and mining products[B]	1.839 (0.8324)	0.395 (0.9724)	0.196 (0.9962)	0.836
Industrial products[C]	1.766 (0.6320)	0.523 (0.9986)	0.034 (0.9983)	0.968
Construction[D]	1.732 (0.6731)	0.524 (0.9986)	0.028 (0.9999)	0.973
Market Services[E]	2.052 (0.5326)	0.603 (0.9731)	0.011 (0.9983)	0.989
Non Market Services[F]	1.393 (0.9415)	0.194 (0.7458)	0.206 (0.9996)	0.829

A: α [0;10 times the calibrated value]; δ [0;1] ρ[-10;10]; [-4.098;4.098] for the error component
B: α [0;10 times the calibrated value]; δ [0;1] ρ[-2.5;2.5]; [-0.667;0.667] for the error component
C: α [0;10 times the calibrated value]; δ [0;1] ρ[-10;10]; [-3.819;3.819] for the error component
D: α [0;10 times the calibrated value]; δ [0;1] ρ[-2.5;2.5]; [-1;1] for the error component
E: α [0;10 times the calibrated value]; δ [0;1] ρ[-5;5]; [-1.798;1.798] for the error component
F: α [0;10 times the calibrated value]; δ [0;1] ρ[-2.5;2.5]; [-1.126;1.126] for the error component
Source: our elaboration on the RSAM for Sardinia (2001)

Appendix 2

Table A2.1 - Expenditure elasticity (GME estimates Household group A)[1]

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.3306	1.3306	1.3306	1.3289	1.1256
Energy and Mining products	3.5650	3.5650	3.5649	3.5535	1.9749
Industrial products	0.9614	0.9614	0.9614	0.9614	0.9853
Market services	1.0034	1.0034	1.0034	1.0035	1.0013
Non Market services	0.9275	0.9275	0.9275	0.9280	0.9725
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98744	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.2 – Expenditure elasticity (GME estimates Household group B)

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.3586	1.3586	1.3586	1.3569	1.1466
Energy and Mining products	2.6207	2.6207	2.6207	2.6142	1.6626
Industrial products	0.9626	0.9626	0.9626	0.9627	0.9847
Market services	1.0024	1.0024	1.0024	1.0024	1.0001
Non Market services	0.9587	0.9587	0.9587	0.9590	0.9831
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98744	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

[1] All the tables reported in this Appendix, refer to the following parameter supports:
Parameter supports:
GME 1 [-100;100] for α and β; [-1;1]for the error terms;
GME 2 [-50;50] for α and β; [-1;1]for the error terms;
GME 3 [-10;10] for α and β; [-1;1]for the error terms;
GME 4 [-1;1] for α and β; [-1;1]for the error terms;
GME 5 [-10;10] for the α and [-1;1] for the β; [-1;1]for the error terms;

Table A2.3 - Expenditure elasticity (GME estimates Household group C)

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.4434	1.4434	1.4434	1.4412	1.1714
Energy and Mining products	2.9854	2.9854	2.9853	2.9765	1.7676
Industrial products	0.9591	0.9591	0.9591	0.9592	0.9842
Market services	1.0047	1.0047	1.0047	1.0047	1.0018
Non Market services	0.9160	0.9160	0.9160	0.9166	0.9676
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98744	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.4 - Expenditure elasticity (GME estimates Household group D)

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.4151	1.4150	1.4150	1.4133	1.1805
Energy and Mining products	3.0499	3.0508	3.0492	3.0417	1.8916
Industrial products	0.9632	0.9632	0.9632	0.9633	0.9840
Market services	1.0044	1.0044	1.0044	1.0044	1.0019
Non Market services	0.9190	0.9190	0.9190	0.9195	0.9648
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98744	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.5 - Expenditure elasticity (GME estimates Household group E)

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.5891	1.5891	1.5891	1.5859	1.2107
Energy and Mining products	2.4519	2.4518	2.4518	2.4443	1.5194
Industrial products	0.9522	0.9522	0.9522	0.9523	0.9829
Market services	1.0066	1.0066	1.0066	1.0066	1.0024
Non Market services	0.9067	0.9067	0.9067	0.9075	0.9667
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98743	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.6 - Expenditure elasticity (GME estimates for Household group F)

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	1.5268	1.5268	1.5267	1.5240	1.1970
Energy and Mining products	2.7768	2.7768	2.7765	2.7680	1.6645
Industrial products	0.9570	0.9570	0.9570	0.9571	0.9839
Market services	1.0055	1.0055	1.0055	1.0055	1.0020
Non Market services	0.8963	0.8964	0.8964	0.8972	0.9613
Normalized Entropy Ratio	0.99999	0.99999	0.99987	0.98743	0.99986

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.7 - Sensitivity Analysis of GME estimates for the β parameters – Household A

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0094 (0.9999)	0.0094 (0.9999)	0.0094 (0.9997)	0.0094 (0.9746)	0.0036 (0.9996)
Energy and Mining products	0.0078 (0.9999)	0.0078 (0.9999)	0.0078 (0.9997)	0.0078 (0.9747)	0.0030 (0.9996)
Industrial products	-0.0131 (0.9999)	-0.0131 (0.9999)	-0.0131 (0.9999)	-0.0131 (0.9753)	-0.0050 (0.9999)
Market services	0.0019 (0.9999)	0.0019 (0.9999)	0.0019 (0.9997)	0.0019 (0.9749)	0.0007 (0.9997)
Non Market services	-0.0060 (0.9999)	0.1992 (0.9999)	-0.0060 (0.9997)	-0.0059 (0.9751)	-0.0023 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.8 - Sensitivity Analysis of GME estimates for the α parameters – Household A

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2012 (0.9999)	0.2012 (0.9999)	0.2012 (0.9999)	0.2011 (0.9999)	0.2462 (0.9999)
Energy and Mining products	0.2010 (1.000)	0.2010 (0.9999)	0.2010 (0.9999)	0.2009 (0.9999)	0.2383 (0.9999)
Industrial products	0.1983 (0.9999)	0.1983 (0.9999)	0.1983 (0.9999)	0.1984 (0.9999)	0.1355 (0.9999)
Market services	0.2002 (1.000)	0.2002 (1.000)	0.2002 (0.9999)	0.2002 (0.9999)	0.2093 (0.9999)
Non Market services	0.1993 (1.000)	0.1992 (0.9999)	0.1992 (0.9999)	0.1993 (0.9999)	0.1707 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.9 - Sensitivity Analysis of GME estimates for the β parameters – Household B

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0085 (0.9999)	0.0085 (0.9999)	0.0085 (0.9997)	0.0085 (0.9747)	0.0035 (0.9996)
Energy and Mining products	0.0074 (0.9999)	0.0074 (0.9999)	0.0074 (0.9997)	0.0073 (0.9747)	0.0030 (0.9996)
Industrial products	-0.0124 (0.9999)	-0.0124 (0.9999)	-0.0124 (0.9997)	-0.0124 (0.9753)	-0.0051 (0.9998)
Market services	0.0012 (0.9999)	0.0012 (0.9999)	0.0012 (0.9997)	0.0012 (0.9749)	0.0005 (0.9997)
Non Market services	-0.0047 (0.9999)	-0.0047 (0.9999)	-0.0047 (0.9997)	-0.0047 (0.9751)	-0.0019 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.10 - Sensitivity Analysis of GME estimates for the α parameters – Household B

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2010 (1.0000)	0.2010 (0.9999)	0.2010 (0.9999)	0.2010 (0.9999)	0.2424 (0.9999)
Energy and Mining products	0.2009 (1.0000)	0.2009 (0.9999)	0.2009 (0.9999)	0.2008 (0.9999)	0.2367 (0.9999)
Industrial products	0.1985 (0.9999)	0.1985 (0.9999)	0.1985 (0.9999)	0.1986 (0.9999)	0.1381 (0.9999)
Market services	0.2001 (1.0000)	0.2001 (1.0000)	0.2001 (0.9999)	0.2001 (0.9999)	0.2062 (0.9999)
Non Market services	0.1994 (1.0000)	0.1994 (0.9999)	0.1994 (0.9999)	0.1995 (0.9999)	0.1766 (0.9999)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.11 - Sensitivity Analysis of GME estimates for the β parameters – Household C

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0088 (0.9999)	0.0088 (0.9999)	0.0088 (0.9997)	0.0087 (0.9746)	0.0034 (0.9996)
Energy and Mining products	0.0078 (0.9999)	0.0078 (0.9999)	0.0078 (0.9997)	0.0077 (0.9747)	0.0030 (0.9996)
Industrial products	-0.0132 (0.9999)	-0.0131 (0.9999)	-0.0132 (0.9997)	-0.0131 (0.9753)	-0.0051 (0.9999)
Market services	0.0027 (0.9999)	0.0027 (0.9999)	0.0027 (0.9997)	0.0027 (0.9748)	0.0011 (0.9997)
Non Market services	-0.0061 (0.9999)	-0.0061 (0.9999)	-0.0061 (0.9997)	-0.0061 (0.9751)	-0.0024 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.12 - Sensitivity Analysis of GME estimates for the α parameters – Household C

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2011 (1.0000)	0.2011 (0.9999)	0.2011 (0.9999)	0.2010 (0.9999)	0.2432 (0.9999)
Energy and Mining products	0.2010 (1.0000)	0.2010 (0.9999)	0.2010 (0.9999)	0.2009 (0.9999)	0.2383 (0.9999)
Industrial products	0.1983 (0.9999)	0.1983 (0.9999)	0.1983 (0.9999)	0.1984 (0.9999)	0.1351 (0.9998)
Market services	0.2003 (1.0000)	0.2003 (1.0000)	0.2003 (0.9999)	0.2003 (0.9999)	0.2136 (0.9999)
Non Market services	0.1992 (1.0000)	0.1992 (0.9999)	0.1992 (0.9999)	0.1993 (0.9999)	0.1697 (0.9999)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.13 - Sensitivity Analysis of GME estimates for the β parameters – Household D

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0079 (0.9999)	0.0079 (0.9999)	0.0079 (0.9997)	0.0079 (0.9747)	0.0034 (0.9996)
Energy and Mining products	0.0070 (0.9999)	0.0070 (0.9999)	0.0070 (0.9997)	0.0070 (0.9747)	0.0031 (0.9996)
Industrial products	-0.0119 (0.9999)	-0.0119 (0.9999)	-0.0119 (0.9997)	-0.0119 (0.9752)	-0.0052 (0.9999)
Market services	0.0026 (0.9999)	0.0026 (0.9999)	0.0026 (0.9997)	0.0026 (0.9748)	0.0011 (0.9997)
Non Market services	-0.0056 (0.9999)	-0.0057 (0.9999)	-0.0056 (0.9997)	-0.0056 (0.9751)	-0.0024 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.14 - Sensitivity Analysis of GME estimates for the α parameters – Household D

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2009 (1.0000)	0.2009 (0.9999)	0.2009 (0.9999)	0.2008 (0.9999)	0.2397 (0.9999)
Energy and Mining products	0.2008 (1.0000)	0.2008 (0.9999)	0.2008 (0.9999)	0.2008 (0.9999)	0.2353 (0.9999)
Industrial products	0.1986 (0.9999)	0.1986 (0.9999)	0.1986 (0.9999)	0.1987 (0.9999)	0.1404 (0.9999)
Market services	0.2003 (1.0000)	0.2003 (1.0000)	0.2003 (0.9999)	0.2003 (0.9999)	0.2129 (0.9999)
Non Market services	0.1993 (1.0000)	0.1994 (0.9999)	0.1993 (0.9999)	0.1994 (0.9999)	0.1717 (0.9999)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.15 - Sensitivity Analysis of GME estimates for the β parameters – Household E

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0087 (0.9999)	0.0087 (0.9999)	0.0087 (0.9997)	0.0087 (0.9746)	0.0031 (0.9996)
Energy and Mining products	0.0081 (0.9999)	0.0081 (0.9999)	0.0081 (0.9997)	0.0081 (0.9747)	0.0029 (0.9996)
Industrial products	-0.0144 (0.9999)	-0.0144 (0.9999)	-0.0144 (0.9997)	-0.0143 (0.9754)	-0.0051 (0.9999)
Market services	0.0042 (0.9999)	0.0042 (0.9999)	0.0042 (0.9997)	0.0042 (0.9748)	0.0015 (0.9997)
Non Market services	-0.0067 (0.9999)	-0.0067 (0.9999)	-0.0067 (0.9997)	-0.0067 (0.9751)	-0.0024 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.16 - Sensitivity Analysis of GME estimates for the α parameters – Household E

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2012 (1.0000)	0.2012 (0.9999)	0.2012 (0.9999)	0.2011 (0.9999)	0.2424 (0.9999)
Energy and Mining products	0.2011 (1.0000)	0.2011 (0.9999)	0.2011 (0.9999)	0.2010 (0.9999)	0.2394 (0.9999)
Industrial products	0.1980 (0.9999)	0.1980 (0.9999)	0.1980 (0.9999)	0.1981 (0.9999)	0.1302 (0.9999)
Market services	0.2006 (1.0000)	0.2006 (1.0000)	0.2006 (0.9999)	0.2005 (0.9999)	0.2204 (0.9999)
Non Market services	0.1991 (1.0000)	0.1991 (0.9999)	0.1991 (0.9999)	0.1991 (0.9999)	0.1674 (0.9999)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.17 - Sensitivity Analysis of GME estimates for the β parameters – Household F

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.0088 (0.9999)	0.0088 (0.9999)	0.0088 (0.9997)	0.0088 (0.9746)	0.0033 (0.9996)
Energy and Mining products	0.0081 (0.9999)	0.0080 (0.9999)	0.0080 (0.9997)	0.0080 (0.9747)	0.0030 (0.9996)
Industrial products	-0.0136 (0.9999)	-0.0136 (0.9999)	-0.0136 (0.9997)	-0.0136 (0.9753)	-0.0051 (0.9999)
Market services	0.0033 (0.9999)	0.0033 (0.9999)	0.0033 (0.9997)	0.0033 (0.9748)	0.0012 (0.9997)
Non Market services	-0.0066 (0.9999)	-0.0066 (0.9999)	-0.0066 (0.9997)	-0.0065 (0.9751)	-0.0025 (0.9998)

Source: our elaboration on the RSAM for Sardinia (2001)

Table A2.18 - Sensitivity Analysis of GME estimates for the α parameters – Household F

	GME 1	GME 2	GME 3	GME 4	GME 5
Agriculture, animal husbandry and fishing	0.2011 (1.0000)	0.2011 (0.9999)	0.2011 (0.9999)	0.2011 (0.9999)	0.2433 (0.9999)
Energy and Mining products	0.2010 (1.0000)	0.2010 (0.9999)	0.2010 (0.9999)	0.2010 (0.9999)	0.2494 (0.9999)
Industrial products	0.1982 (0.9999)	0.1982 (0.9999)	0.1982 (0.9999)	0.1983 (0.9999)	0.1334 (0.9999)
Market services	0.2004 (1.0000)	0.2004 (1.0000)	0.2004 (0.9999)	0.2004 (0.9999)	0.2162 (0.9999)
Non Market services	0.1992 (1.0000)	0.1991 (0.9999)	0.1991 (0.9999)	0.1992 (0.9999)	0.1677 (0.9999)

Source: our elaboration on the RSAM for Sardinia (2001)

Bibliography

Abdelkhalek T. and Dufour J.M. (1998), *Statistical Inference for Computable General Equilibrium Models, with application to a model of the Moroccan Economy*, The Review of Economics and Statistics, 80 (4), 520-534.

Adams R.H. Jr (2005), Remittances, *Household Expenditure and Investment in Guatemala*, World Bank Policy Research, Working Paper 3532.

Adkins L.C., Rickman D.S. and Hameed A. (2003), *Bayesian Estimation of Regional Production for CGE Modeling*, Journal of Regional Science, 43 (4), 641-661.

Akbay C., Boz I. and Chern W.S. (2007), *Household Food Consumption in Turkey,* European Review of Agricultural Economics, 34 (2), 209–231.

Alaouze C.M. (1977), *Estimates of the Elasticity of Substitution between Imported and Domestically Produced Goods Classified at the Input-Output Level of Aggregation,* IMPACT Project Paper O-13.

André F., Cardenete M. and Velázquez E. (2005), *Performing an Environmental Tax Reform in a Regional Economy. A Computable General Equilibrium Approach,* The Annals of Regional Science, Springer, 39 (2), 375-392.

Annabi N., Cockburn J. and Decaluwé B. (2006), *Functional Form and Parametrization of CGE Models,* MPIA Working Paper 2006-04, Poverty and Economic Policy (PEP) Research Network, International Development Research Centre (IDRC).

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.

Arndt C., Cruz A., Jensen H.T., Robinson S. and Tarp F. (1998), *Social Accounting Matrices for Mozambique: 1994-95,* International Food Policy Research Institute, Trade and Macroeconomics Division Paper No. 28.

Arrow KJ, Chenery H.B., Minahs B.S. and Solow R.M (1961), *Capital-Labour Substitution and Economic Efficiency,* Review of Economics and Statistics, 43 (3), 225-247.

Balcombe K, Davidova S. and Morrison J.A. (1999), *Consumer Behaviour in a Country in Transition with a Strongly Contracting Economy: The Case of Food Consumption in Bulgaria,* Journal of Agricultural Economics, 50 (1), 36-47.

Balcombe K, Rapsomanikis G. and Klonaris S. (2003), *Cross Entropy Estimation of the Almost Ideal Demand System for Greek Consumption,* Journal of Agricultural Economics, 54 (3), 447-466.

Balistreri E.J and Hillberry R.H. (2005), *Estibration: an Illustration of Structural Estimation as Calibration,* Colorado School of Mines- Division of Economics and Business and University of Melbourne.

Balistreri, E.J., McDaniel C.A. and Wong E.V. (2002), *An Estimation of US Industry- Level Capital Labour Substitution Elasticities: Cobb-Douglas as a Reasonable Starting Point?,* Working Paper.

Bandara J.S. (1991), *Computable General Equilibrium Models for Development Policy Analysis in LDCs,* Journal of Economic Surveys, 7, 3-69.

Banks J., Blundell R. and Lewbel A. (1997), *Quadratic Engel Curves and Consumer Demand,* Review of Economics and Statistics, 79 (4), 527-539.

Beneito P. (2003), *A Complete System of Engel Curves in the Spanish Economy,* Applied Economics, 35 (7), 803-816.

Bera A.K. and Bilias Y. (2002), *The MM, ME, ML, EL, EF and GMM Approaches to Estimation: a Synthesis,* Journal of Econometrics, 107, 51-86.

Bergman L. (1982), *A System of Computable General Equilibrium Models for a Small Open Economy,* Mathematical Modelling, 3, 421-435.

Bergman, L., and Henrekson, M. (2003), *CGE Modeling of Environmental Policy and Resource Management*, presented at 2nd Advanced Course on Computable General Equilibrium Modelling (CGE) and the Environment, Trieste, December 1-12, 2003.

Bhattacharyya S.C. (1996), *Applied General Equilibrium Models for Energy Studies: A Survey,* Energy Economics, 18, 145-164.

Blackorby C. and Russell R.R. (1989), *Will the Real Elasticity of Substitution Please Stand Up? A Comparison of the Allen/Uzawa and Morishima Elasticities,* American Economic Review, 79, 882-88.

Blundell R. and Duncan A. (1998), *Kernel Regression in Empirical Microeconomics,* Journal of Human Resources, 33 (1), 62-87

Boys A.K. and Florax R.J.G.M. (2007), *Meta-Regression Estimates for CGE Models: a Case Study for Input Substitution Elasticities in Production Agriculture,* Preliminary version, Department of Agricultural Economics, Purdue University Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Portland, OR, July 29 – August 1, 2007.

Britz W. and Wieck C. (2002), *Completeness and Consistency in a Multidimensional Data Base using Constrained Simultaneous Estimation Techniques,* Poster presented at the 42 Jahrestagung der GEWISOLA (German Agricultural Economics Society).

Browing M, Hansen L.P. and Heckman J.J. (1999), *Micro Data and General Equilibrium Models,* in Taylor J. and Woodford M., eds., Handbook of Macroeconomics, Vol. 1A (North-Holland, Amsterdam), Chapter 8, pp. 543-633.

Browne M., Ortmann G.F. and Hendriks S. (2007), *Expenditure Elasticities for Rural Households in the Embo ward,* Umbumbulum, KwaZulu-Natal, Agrekon 46 (4).

Bruno M. (1984), *Raw Materials, Profits and the Productivity Slowdown,* The Quarterly Journal of Economics, 99 (1), 1-30.

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

Castaldo A. and Reilly B. (2007), *Do Migrant Remittances Affect the Consumption patterns of Albanian Households?*, South-Eastern Europe Journal of Economics, 1, 25-54.

Centeno Stenberg L. and Siriwardana M. (2005), *The Appropriateness of CGE Modelling in Analysing the Problem of Deforestation,* Management of Environmental Quality: An International Journal, 16 (5), 407-420.

Chambers R.G. (1988), *Applied Production Analysis. A Dual Approach,* University of Maryland, Cambridge University Press.

Chern W.S, Ishibashi K., Taniguchi K. and Tokoyama Y. (2003), *Analysis of Food Consumption Behaviour by Japanese Households,* Fao Economic and Social Development Paper n. 152.

Chiang A.C. (1984), *Fundamental Methods of Mathematical Economics,* McGraw-Hill international editions, Economic series.

Clarke B. (2007), *Information optimality and Bayesian Modelling,* Journal of Econometrics, 138, 405-429.

Clemhout S. (1968), *The Class of Homothetic Isoquant Production Functions,* The review of Economic studies, 35 (1), 91-104.

Dawkins C., Srinivasan T.N. and Whalley J. (2001), *Calibration,* in James J. Heckman and Edward Leamer (eds.), Handbook of Econometrics. Vol. 5, pp.3653-3703, Amsterdam, the Netherlands: Elservier Science.

De Boer P. and Missaglia M. (2006), *Estimation of income elasticities and their use in a CGE model for Palestine,* Econometric Institute Report EI 2006-12.

De Mello M, Pack A and Sinclair M.T. (2002), *A System of Equations Model of UK Tourism Demand in Neighbouring Countries*, Applied Economics, 34, 509-521.

De Melo J. (1988), *Computable General Equilibrium Models for Trade Policy Analysis in Developing Countries: A Survey,* Journal of Policy Modeling, 10, 469-503.

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

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

Deaton A. and Muellbauer J. (1980b), Economics and Consumer Behavior, Cambridge University Press, London.

Decaluwè B. and Martens A. (1988), *CGE Modeling and Developing Economies: A Concise Empirical Survey of 73 Application to 26 Countries,* Journal of Policy Modeling, 10 (4), 529-668.

Devarajan S. and Robinson S. (2002), *The Influence of Computable General Equilibrium Models on Policy,* TMD discussion paper no.98, International Food Policy Research Institute, Washington DC., USA.

Diewert W.E. and Lawrence D. (1994), *The Marginal Costs of Taxation in New Zealand,* Report prepared for the New Zealand Business Round Table by Swan Consultants Pty Ltd, Canberra.

Dixon P.B, Parmenter B.R., Sutton J. and Vincent D.P. (1982), *Orani: a Multisectoral Model of the Australian Economy,* North Holland, Amsterdam.

Dixon P.B. (2006), *Evidence-Based Trade Policy Decision Making in Australia and the Development of Computable General Equilibrium Modelling,* Centre of Policy Studies, IMPACT Centre Working Papers, g-163, Monash University.

Eruygur H.O. (2005), *Generalized Maximum Entropy Estimator: Formulation and a Monte Carlo Study,* Paper presented at the VII National Symposium on Econometrics and Statistics Istanbul, Turkey, May 26-27, 2005.

Estrin S. and Laidler D. (1995), *Introduction to Microeconomics* (fourth Edition), Harvester Wheatsheaf, New York.

Ferrari G. (1998), *Alla ricerca della Matrice Perduta,* in '7me Colloque de Comptabilité Nationale, Paris 28-30 Janvier 1998', pp. 1-17.

Ferrari G. (1999), *Introduzione ai sistemi di contabilità nazionale,* Centro 2P edizioni, Firenze.

Ferrari G. and Manca A. (2008), *Estimating the Parameters of a CES Production Function in a Regional Environmentally Extended CGE Model Framework: a RESAM Only Based GME Approach,* Dipartimento di Statistica "G.Parenti, Università degli Studi di Firenze.

Ferrari G., Garau G. and Lecca P. (2007), *A RSAM for Sardinia,* mimeo, Dipartimento di Statistica, Università di Firenze.

Fraser I. (2000), *An Application of Maximum Entropy Estimation: the Demand for Meat in the United Kingdom,* Applied Economics, 32 (1), 45-59.

Garau G. and Lecca P. (2008), *Impact Analysis of Regional Knowledge Subsidy: a CGE Approach,* CRENOS Working papers, no 2008/11.

Gebreselasie T.G. (2008), *Sectoral Elasticity of Substitution and Returns to Scale in South Africa,* South African Journal of Economics, vol. 76:S2.

Giusti F. (1994), *Modelli neoclassici di produzione – evoluzione storica e ricerca quantitative,* Università degli studi di Roma "La Sapienza", Dipartimento di Teoria Economica e metodi quantitativi per le scelte politiche.

Golan A. (2001), *A Simultaneous Estimation and Variable Selection Rule,* Journal of Econometrics 101 (1), 165-193.

Golan A. (2002), *Information and Entropy Econometrics – Editor's view,* Journal of Econometrics 107 (1-2), 1-15.

Golan A. (2003), *Image Reconstruction of Incomplete and Ill-posed Data: An Information-Theoretic Approach,* Technical report, Department of Economics, American University.

Golan A. (2007), *Information and Entropy Econometrics – Volume Overview and Synthesis,* Journal of Econometrics, 138, 379-387.

Golan A. (2008) Information and entropy econometrics – A review and synthesis, Foundations and Trends in Econometrics, Vol. 2 Nos. 1-2 (2006), 1:145.

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. and Gzyl H. (2003), *Priors and Information-Theoretic Estimation,* American Statistical Association Proceedings.

Golan A. and Gzyl H. (2006), *Ill Conditioned Linear Estimation Problems: an Information Theoretic Approach with Priors,* Working Paper.

Golan A. and Perloff J. (2002), Comparison of maximum entropy and higher-order entropy estimators, Journal of Econometrics, 107 (1-2), pp. 195-211.

Golan A. and Vogel S. (2000), *Estimation of Non Stationary Social Accounting Matrix Coefficients with Supplied Side Information,* Economic System Research, 12 (4), 447-471.

Golan A., Judge G. and Miller D. (1996a), *Maximum Entropy Econometrics: Robust Estimation with Limited Data,* John Wiley and Sons, New York.

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

Golan A., Judge G.G. and Karp L. (1996c), *A Maximum Entropy Approach to Estimation and Inference in Dynamic Models or Counting Fish in the Sea Using Maximum Entropy,* Journal of Economic Dynamics and Control, 20, 559-582.

Golan A., Karp L., Perloff J.M. (1996d), *Estimating a Mixed Strategy Employing Maximum Entropy,* CUDARE working papers n. 780, Department of Agricultural and Resource Economics, University of California, Berkeley.

Golan A., Perloff J.M. and Shen E.Z. (2001), *Estimating a Demand System with Nonnegativity Constraints: Mexican Meat Demand,* The Review of economics and statistics, 83 (3), 541-550.

Good I.J. (1963), *Maximum Entropy for Hypothesis Formulation,* Especially for Multidimensional Contingency Tables, Annals of Mathematical Statistics, 34, 911-934.

Gradzewicz M., Griffin P. and Zòlkiewski Z. (2006), *An Empirical Recursive-Dynamic General Equilibrium Model of Poland's Economy. Including Simulations of the Labor Market Effects of Key Structural Fiscal Policy Reforms,* Document of the World Bank and National Bank of Poland, Warsaw.

Guarini R. and Tassinari F. (2002), *Statistica economica. Problemi e metodi di analisi,* Il Mulino, Bologna.

Hansen L.P and Heckman J.J. (1996), *The Empirical Foundations of Calibration*, The Journal of Economic Perspectives, 10 (1), 87-104.

Heckelei T. and Wolff H. (2003), *Estimation of Constrained Optimisation Models for Agricultural Supply Analysis Based on Generalised Maximum Entropy,* European Review of Agricultural Economics, 30 (1), 27-50.

Heckelei T., Mittelhammer R. and Jansson T. (2008), *A Bayesian Alternative to Generalized Cross Entropy Solutions for Underdetermined Econometric Models,* discussion paper no. 2/2008, Institute for food and resource economics, University of Bonn.

Hertel T, Hummels D. Ivanic M. and Keeney R. (2007), *How confident can we be of CGE-based assessments of Free Trade Agreements?*, Economic Modelling, 24,611–635.

Hickman B. and Lau L. (1973), *Elasticity of Substitution in a World Trade Model,* European Economic Review 4, 347-180.

Hicks J.R. (1932), *Theory of Wages,* London, MacMillan.

Hoff A. (2002), *The Translog Approximation of the Constant Elasticity of Substitution Production Function with more than Two Input Variables,* working paper 14/2002, Danish research institute of food economics.

Hoff A. (2004), *The Linear Approximation of the CES Function with n Input Variables,* Marine resource economics, 19, 295-306.

Holste D., Grobe I. and Herzel H. (1998), *Bayes' Estimators of Generalized Entropies,* J. Phys. A: Math. Gen., 31, 2551-2566.

Hoover K.D. (1995), *Facts and Artifacts: Calibration and the Empirical Assessment of Real Business Cycle Models,* Oxford Economic Papers, 45, 24-45.

Howitt R.E. and Msangi S. (2006), *Estimating Disaggregate Production Functions: an Application to Northern Mexico,* paper presented at the Annual Meeting of the American Agricultural Economics Association, Long Beach, California, July 23-26.

Howitt R.E. and Reynaud A. (2003), *Spatial Disaggregation of Agricultural Production Data using Maximum Entropy,* European Review of Agricultural Economics, 30, 359-387.

Iqbal Z. and Siddiqui R., (2001), *Critical Review of Literature on Computable General Equilibrium Models,* Pakistan Institute of Development Economics, Mimap Technical Paper Series, n. 9, Islamabad, Pakistan.

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

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

Jaynes E.T. (1963), *Information Theory and Statistical Mechanics II,* in K.W. Ford (ed.), Statistical Physics, W.A. Benamin, Inc., New York, 181-218.

Jaynes E.T. (1982), *On the Rationale of Maximum-Entropy Methods,* in Proc. IEEE 70, 939-952.

Jaynes E.T. (1984), *Prior Information and Ambiguity in Inverse Problems,* in McLaughlin D.W, ed., "Inverse Problems: proceedings of the Symposium in Applied Mathematics of the American Mathematical Society and the Society for Industrial and Applied Mathematics, held in New York, April 12-13th , American Mathematical Society, Providence, R.I, 151-166.

Johansen L. (1960), *A Multi-Sectoral Study of Economic Growth,* North Holland, Amsterdam.

Jorgensen D.W., Slenick D.T., Wilcoxen P.J. (1992), *Carbon Taxes and Economic Welfare,* Research Discussion Paper 1589, Harvard Institute of Economic.

Jorgenson D.W. (1983), *Modeling Production for General Equilibrium Analysis,* The Scandinavian Journal of Economics, vol. 85, no.2, Topics in Production Theory, pp. 101-112.

Jorgenson D.W. and Fraumeni B.M. (1981), *Relative Prices and Technical Change, in Modeling and Measuring Natural Resource Substitution* (E.Berndt and B. Fields, eds.), Cambridge, MIT Press, pp.17-47. Reprinted in Quantitative Studies on Production and Price, (Eichhorn W., Henn R., Neumann K. and Sheppard R.W., eds.), Wurzburg, Physic-Verlag (1983), pp. 241-69.

Jorgenson D.W. and Laffont J.J. (1974), *Efficient Estimation of Non-linear Simultaneous Equations with Additive Disturbances,* Annals of Economic and Social Measurement, 3 (4), 615-40.

Jorgenson D.W. and Slesnich D.T. (1997), *General Equilibrium Analysis of Economic Policy in Welfare,* Jorgenson D.W., ed., vol.2, Cambridge, The MIT Press.

Jorgenson D.W.(1984), *Econometric Methods for Applied General Equilibrium Analysis,* in J. Shoven and J. Whalley, eds, 'Applying General Equilibrium', Cambridge, Cambridge University Press.

Jorgenson D.W., Lau L.J. and Stocker T.M. (1980), *Welfare Comparison Under Exact Aggregation,* American Economic Review, 70, No. 2 (May), pp. 268-72.

Judge G.G. and Golan A. (1992), *Recovering Information in the Case of Ill-Posed Inverse Problem with Noise.* Mimeo, Department of Agricultural and Natural Resources, University of California, Berkeley, CA.

Kadiyala K.R. (1972), *Production Functions and Elasticity of Substitution,* Southern Economic Journal, 38 (3), 281-284.

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

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

Kesavan H.K. and Kapur J.N (1989), *The Generalized Maximum Entropy Principle,* IEEE Transactions on systems, Man, and Cybernetics,19 (5).

Kmenta J. (1967), *On Estimation of the CES Production Function,* International Economic Review, 8 (2), 180-189.

Kouparitsas M.A. (2001), *Should Trade Barriers be Phased-Out Slowly? A Case Study of North America,* Journal of Policy Modeling, 23, 875-900.

Kullback S. (1959), Information Theory and Statistics, John Wiley & Sons, New York.

Kullback S. and Leibler R.A. (1951), *On Information and Sufficiency,* Ann. Math. Stat., 22, 79-86.

Lansink A.O. (1999), *Generalized Maximum Entropy Estimation and Heterogeneous Technologies,* European Review of Agricultural Economics, 26 (1), 101-115.

Lau L. (1984), *Comments on Mansur and Whalley's Numerical Specification of Applied General Equilibrium Models,* In Scarf. H. and Shoven J.B (Eds) pp. 127-137.

Laureti T. (2006), *L'efficienza rispetto alla frontiera delle possibilità produttive. Modelli teorici ed analisi empiriche,* Firenze University Press.

Lee T.C. and Judge G.G. (1996), *Entropy and Cross-Entropy Procedures for Estimating Transition Probabilities from Aggregate Data,* in Bayesian Analysis in Statistics and Econometrics, edited by D.A. Barry, K.M. Chaloner and J.K. Geweke, New York: John Wiley & Sons.

Lence S.H and Miller D.J (1998), *Estimation of Multi-Output Production Functions with Incomplete Data: a Generalised Maximum Entropy Approach,* European review of agricultural economics, 25, 188-209.

Léon Y., Peeters L., Quinqu M. and Surry Y. (1999), *The Use of Maximum Entropy to Estimate Input-Output Coefficients from Regional Farm Accounting Data,* Journal of Agricultural Economics, 50, 425-439.

Leontief W. (1941), *The Structure of American Economy 1919-1929,* Harvard University Press, Cambridge.

Leser C.E.V. (1963), *Forms of Engel Functions,* Econometrica, vol. 31, no.4, pp. 694: 703.

Liu J., Arndt C. and Hertel T.W. (2003), *Parameter Estimation and Measures of Fit in A Global General Equilibrium Model*, GTAP Working Paper No.23, West Lafayette IN, Purdue University.

Lofgren H., Lee Harris R. and Robinson S. (2002), *A Standard Computable General Equilibrium (CGE). Model in GAMS,* International Food Policy Research Institute.

Manca A. (2007), *Calibrazione e stima di un modello di equilibrio economico regionale esteso all'ambiente per la Sardegna,* Tesi di Dottorato, Dipartimento di Statistica "G.Parenti, Università degli Studi di Firenze.

Mansur A. and Whalley J (1984), *Numerical Specification of Applied General Equilibrium Models: Estimation, Calibration and Data,* in J.Shoven & J.Whalley, eds, Applying General Equilibrium, Cambridge: Cambridge University Press, pp. 69-127.

McDaniel C.A and Balistreri E.J. (2002), *A Review of Armington Trade Substitution Elasticities,* Research Division, Economics Office, U.S. International Trade Commission.

McElroy F.W.(1967), *Note on the CES Production Function,* Econometrica, 35, 1, 154-156.

McKitrick, R.R. (1998), *The Econometric Critique of Computable General Equilibrium Modeling: the Role of Functional Forms,* Economic Modelling 15, 543–573.

Miller E. (2008), *An Assessment of CES and Cobb-Douglas Production Functions,* working paper no. 05/2008, Congressional Budget Office.

Mishra S.K., (2008), *A Brief History of Production Functions,* working paper series n. 1020577, Department of Economics North-Eastern Hill University, Shillong (India).

Mittelhammer R.C. and Cardell N.S. (1996), *On the Consistency and Asymptotic Normality of the Data-Constrained GME Estimator of the GLM,* mimeo, Pullman, WA: Washington State University.

Mittelhammer R.C. Judge G.G and Miller D.J. (2000), *Econometric Foundations,* Cambridge University Press.

Miyagiwa K. and Papageorgiou C. (2007), *Endogenous Aggregate Elasticity of Substitution,* Journal of Economic Dynamics and Control, 31, 2899-2919.

Muellbauer J. (1976), *Community Preferences and the Representative Consumer,* Econometrica 44, 979-99.

Myung J., Ramamoorti S. and Bailey D. (1996), *Maximum Entropy Aggregation of Expert Predictions,* Management Science, 42 (10), 1420-1436.

Nganou Nguessa J.P. (2004), *Estimating the Key Parameters of the Lesotho CGE Model,* Paper prepared for the International Conference "I/O and GE: data, modeling and policy analysis, Brussels (Belgium), September 2-4, 2004.

Pagan A.R. and Shannon J.H. (1985), *Sensitivity Analysis for Linearized Computable General Equilibrium Models,* dans Piggot J.R. amd Walley J. (eds.), New Development in Applied General Equilibrium Analysis, Cambridge University Press.

Paris Q. and Caputo M.R. (2001), *Sensitivity of the GME Estimates to Support Bounds,* Working Paper No. 01-008, Department of Agricultural and Resource Economics, University of California Davis.

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.

Partridge M. and Rickman D. (1998), *Regional Computable General Equilibrium Modeling: a Survey and Critical Appraisal,* International Regional Science Review, 21 (3), 205-248.

Partridge M. and Rickman D. (2008), *Computable General Equilibrium (CGE) Modelling for Regional Economic Development Analysis*, Regional Studies, 1-18.

Perloff. J., Karp L. and Golan A (2007), *Estimating Market Power and Strategies,* Cambridge University Press.

Piggott J. and Whalley J. (1992), *New Development in Applied General Equilibrium Analysis,* Cambridge University Press, Cambridge.

Piggott J.R. and Whalley J. (1980), *Economic Effects of U.K. Tax-Subsidy Policies: A General Equilibrium Appraisal,* Mimeo.

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

Pyatt G. and Round J. (1985), *Social Accounting Matrices: a Basis for Planning,* Washington D.C., World Bank.

Pyatt, G. (1999), *Some Relationships between T-Accounts, Input-Output Tables and Social Accounting Matrices*, Economic Systems Research, 11 (4), 365 – 387.

Raa T.T (2007), *The Extraction of Technical Coefficients from Input and Output Data,* Economic system research, 19 (4), 453:459.

Rényi A. (1961), *On Measures of Entropy and Information,* in Proceeding of the 4th Berkeley Symoposium on Mathematical Statistics and Probability, Vol. I Berkeley: University of California Press, pp. 547-561.

Rényi A. (1970), *Probability Theory,* Amsterdam: North-Holland.

Rezek J. P. and Campbell R.C. (2007), *Cost Estimates for Multiple Pollutants: a Maximum Entropy Approach,* Energy Economics, 29, 503-519.

Robilliard A.S. and Robinson S. (2003), *Reconciling Household Surveys and National Accounts Data using a Cross Entropy Estimation Method,* Review of Income and Wealth, 49, 395-406.

Robinson S. (2003), *Macro Models and Multipliers: Leontief, Stone, Keynes, and CGE Models,* International Food Policy Research Institute (IFPRI).

Robinson S., Cattaneo A. and El-Said M. (2000), *Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods,* Economic System Research, 13, 47-64.

Roland Holst D.W., Reinert K.A. and Shiells C.R. (1994), *A General Equilibrium Analysis of North American Economic Integration,* in Modeling Trade Policy: Applied General Equilibrium Assessments of North American Free Trade, ed. J.F. Francois and C.R. Shiells, New York, Cambridge University Press, pp. 47-82.

Rosenthal R.E. (tutorial by) (2008), *GAMS – A User's Guide*, GAMS Development Corporation, Washington D.C., USA.

Rutherford T.F. (1999), *Applied General Equilibrium Modeling with MPSGE as a GAMS Subsystem: An Overview of the Modeling Framework and Syntax,* Computational Economics, 14, 1-46.

Ryan T.M. (1973), *CES Production Functions in British Manufacturing Industry: a Cross-Section Study,* Oxford Economic papers, new series, 25 (2), 241-250.

Salem H.H (2001), *The Estimation of the Elasticity of Substitution of a CES Production Function: case of Tunisia,* GAINS, University of Maine.

Sánchez C.M.V. (2004), *Rising inequality and falling poverty in Costa Rica's agriculture during trade reform. A macro-micro general equilibrium analysis*, Maastricht, Shaker.

Sánchez M. and Ros V. (2007), *Informative Note on Elasticities and Calibration of MAMS*, Technical Report prepared for the project Assessing Development Strategies to Achieve the Millenium Development Goals in the Arab Region, Department of Economics and Social Affairs, United Nations.

Saveyn D. and Regemorter Van D. (2007), *Environmental Policy in a Federal State. A Regional CGE Analysis of the NEC Directive in Belgium,* Working Paper series n.2007/01*,* Katholieke Universiteit Leuven, Faculty of Economics and Applied Economic Sciences, Center for economic studies Energy, Transport and Environment.

Scarf. H (1967), *An Example of Algorithm for Calculating General Equilibrium Prices,* The American Economic Review, 59 (4), 669-677.

Schmalensee R., Stocker T.M. and Judson R.A. (1998), *World Carbon Dioxide Emission: 1950-2050,* The Review of Economics and Statistics, 80, 15-27.

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

Shen E.Z. and Perloff J.M. (2001), *Maximum Entropy and Bayesian Approaches to the Ratio Problem,* Journal of Econometrics, 104 (2), 289-313.

Shoven J. and Whalley J. (1982), *General Equilibrium Models for Development Policy,* World Bank Research Publication, Washington D.C..

Shoven J. and Whalley J. (1984), *Applied General Equilibrium Models of Taxation and International Trade: An Introduction and Survey,* Journal of Economic Literature, 22 (3), 1007-1051.

Shoven J. and Whalley J. (1992), *Applying General Equilibrium,* Cambridge: Cambridge University Press.

Singleton, K. (1988). *Econometric Issues in the Analysis of Equilibrium Business Cycle Models*, Journal of Monetary Economics, 21, 361-386.

Sinha A. (2008), *Social Accounting Matrix Analysis with Market and Non-Market Work: A Gender Perspective*, paper presented at VATT, Helsinki, Finland (June).

Skovoroda R.A. (1999), *Estimation of Demand Responses to a Price Change Under the Paucity of the Data,* Working Paper, Mockba.

Sobarzo H.E. (1992), *A General Equilibrium Analysis of the Gains from Trade for the Mexican Economy of a North American Free Trade Agreement,* World Economy, 15 (1), 83-100.

Soofi E.S. (1996), *Information Theory and Bayesian Statistics,* in Berry D.A., Chaloner K.M. and Geweke J.K. (eds): Bayesian Analysis in Statistics and Econometrics: Essays in Honor of Arnold Zellnge, New York: Wiley, pp. 179-189.

Soofi E.S. (2000), *Principal Information Theoretic Approaches,* Journal of the American Statistical Association, 95 (452), 1349-1353.

Stern R.M., Francis J. and Schumacher B. (1976), *Price Elasticities in International Trade: An Annotated Bibliography,* (Macmillan, London, for the Trade Policy Research Centre).

Stone R.(1977), *Social Accounting for Development Planning,* (forward to Pyatt G. and Roe A. et al.) Cambridge University Press, Cambridge, xvi-xxxi.

Stone, R. (1962), *A Social Accounting Matrix for 1960,* Chapman and Hall,Cambridge (UK).

Tassinari G. and Viviani A. (1992), *I consumi delle famiglie italiane: un'analisi mediante sistemi di funzioni di Engel,* Rapporto di ricerca n.10 CON PRI – La misura dei consumi privati, Dipartimento di Scienze Statistiche "Paolo Fortunati", Università degli Studi di Bologna.

Thorbecke, E. (2000), *The Use of Social Accounting Matrices in Modeling,* in 26th General Conference of The International Association for Research in Income and Wealth, Cracow, 27 August-2 September 2000', pp. 1–22.

Tonini A. and Jongeneel R. (2008), *Modelling Dairy Supply for Hungary and Poland by Generalised Maximum Entropy Using Prior Information,* European Review of Agricultural Economics, pp. 1-28.

Tsallis C. (1988), *Possible Generalization of Bolzmann-Gibbs Statistics,* J. Stat. Phys 52, 479-487.

Van Driel H., Nadall V. and Zeelenberg K. (1997), *The Demand for Food in the United States and the Netherlands: A Systems Approach with the CBS Model,* Journal of Applied Econometrics, vol. 12, No. 5, Special Issue: The Experiment in Applied Econometrics (Sep.-Oct., 1997), pp. 509-523, John Wiley & Sons.

Vargas E., Schreiner D., Tembo G. and Marcouiller D. (1999), *Computable General Equilibrium Modeling for Regional Analysis,* Thecnical Report.

Varian H.R. (1992), *Microeconomic Analysis* (third edition), Norton International Student edition, W.W Norton & Company, New York.

Varian H.R. (2002), *Microeconomia,* (edizione italiana a cura di Alfredo Medio),Cafoscarina, Venezia.

Whalley J. (1985), *Hidden Challenges in Recent Applied General Equilibrium Exercises,* in Piggott J. and Whalley J., eds., New Developments in Applied General Equilibrium Analysis (Cambridge University Press, New York).

Wigle R.M. (1986), *General Equilibrium Analysis of the Effects of Multilateral an Canadian Trade Liberalization,* Thesis (Ph.D.), University of Western Ontario.

Wigle R.M. (1991), *The Pagan-Shannon Approximation: Unconditional Systematic Sensitivity in Minutes,* in Piggott J. and Whalley J., eds., Applied General Equilibrium (Physica Verlag, Heidelberg).

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.

Wise J.R., Cooke S. and Holland D. (2002), *A Computable General Equilibrium Analysis of a Property Tax Limitation Initiative in Idaho,* Land Economics, 78 (2).

Witzke H.P. and Britz W. (1998), *A Maximum Entropy Approach to the Calibration of Highly Differentiated Demand Systems,* Working Paper, CAPRI 98-06, University of Bonn, Bonn.

Working H. (1943), *Statistical Laws of Family Expenditure,* Journal of the American Statistical Association, 38 (221), 43-56.

Zellner A. (1988), *Optimal Information Processing and Bayes Theorem,* American Statistician, 42, 278-284.

Zhang X. and Fan S. (2001), *Estimating Crop-Specific Production Technologies in Chinese Agriculture: A Generalized Maximum Entropy Approach,* Am J. Agr. Econ. 83, 378-388.