DE GRUYTER
OPEN

Second Bwanakare

NON-EXTENSIVE ENTROPY ECONOMETRICS FOR LOW FREQUENCY SERIES

NATIONAL ACCOUNTS-BASED INVERSE PROBLEMS

sciendo

Second Bwanakare

Non-Extensive Entropy Econometrics for Low Frequency Series

National Accounts-Based Inverse Problems

Second Bwanakare

Non-Extensive Entropy Econometrics for Low Frequency Series

National Accounts-Based Inverse Problems

Managing Editor: Konrad Sarzyński

Language Editor: Naguib Lallmahomed

DE GRUYTER OPEN

sciendo

Published by De Gruyter Poland Ltd, Warsaw/Berlin
Part of Walter de Gruyter GmbH, Berlin/Boston
The book is published with open access at www.degruyter.com.

Library of Congress Cataloging-in-Publication Data
A CIP catalog record for this book has been applied for at the Library of Congress.

Second Edition

ISBN: 978-3-11-060590-7
e-ISBN: 978-3-11-060591-4

Bibliographic information published by the Deutsche Nationalbibliothek. The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.dnb.de.

Managing Editor: Konrad Sarzyński
Language Editor: Naguib Lallmahomed

www.degruyter.com

Cover illustration: © Thinkstock, Credit: liuzishan

To my wife Rose and my daughter Ozane for their many sacrifices

Contents

Acknowledgements —— xi
Summary —— xii

PART I: **Generalities and Scope of the Book**

1 Generalities —— 2
1.1 Information-Theoretic Maximum Entropy Principle and Inverse Problem —— **2**
1.1.1 Information-Theoretic Maximum Entropy Principle —— **2**
1.2 Motivation of the Work —— **4**
1.2.1 Frequent Limitations of Shannon-Gibbs Maximum Entropy Econometrics —— **4**
1.2.2 Rationale of PL-Related Tsallis Entropy Econometrics and Low Frequency Series —— **6**
1.3 National Accounts-Related Models and the Scope of this Work —— **10**
Bibliography – Part I —— **11**

PART II: **Statistical Theory of Information and Generalised Inverse Problem**

1 Information and its Main Quantitative Properties —— 16
1.1 Definition and Generalities —— **16**
1.2 Main Quantitative Properties of Statistical Information —— **19**
2 Ill-posed Inverse Problem Solution and the Maximum Entropy Principle —— 22
2.1 Introduction —— **22**
2.2 The Inverse Problem and Socio-Economic Phenomena —— **22**
2.2.1 Moore-Penrose Pseudo-Inverse —— **23**
2.2.2 The Gibbs-Shannon Maximum Entropy Principle and the Inverse Problem —— **25**
2.2.3 Kullback-Leibler Cross-Entropy —— **27**
2.3 General Linear Entropy Econometrics —— **28**
2.3.1 Reparametrization of Parameters —— **28**
2.4 Tsallis Entropy and Main Properties —— **29**
2.4.1 Definition and Shannon-Tsallis Entropy Relationships —— **29**
2.4.2 Characterization of Non-Extensive Entropy —— **31**
2.4.2.1 Correlations —— **31**
2.4.2.2 Concavity —— **32**
2.4.3 Tsallis Entropy and Other Forms of Entropy —— **33**

2.4.3.1 Characterization —— **34**
2.4.3.2 Scale of q-Tsallis Index and its Interpretation —— **35**
2.5 Kullback-Leibler-Tsallis Cross-Entropy —— **35**
2.5.1 The q-Generalization of the Kullback-Leibler Relative Entropy —— **35**
2.5.2 Tsallis Versions of the Kullback-Leibler Divergence in Constraining Problems —— **37**
2.6 A Generalized Linear Non-Extensive Entropy Econometric Model —— **38**
2.6.1 A General Model —— **38**
2.6.2 Parameter Confidence Interval Area —— **40**
2.7 An Application Example: a Maximum Tsallis Entropy Econometrics Model for Labour Demand —— **42**
2.7.1 Theoretical Expectation Model —— **42**
2.7.2 A Generalized Non-Extensive Entropy Econometric Model —— **43**
2.7.2.1 General Model —— **43**
2.7.3 Estimated Confidence Area Of Parameters —— **44**
2.7.4 Data and Model Outputs —— **45**
Annex A —— 48
Annex B: Independence of Events Within q-Generalized Kullback-Leibler Relative Entropy —— 49
Bibliography – Part II —— **50**

Part III: **Updating and Forecasting Input-Output Transaction Matrices**

1 Introduction —— 54
2 The System of National Accounts —— 55
3 The Input-Output (IO) Table and its Main Application —— 56
3.1 The I-O Table and Underlying Coefficients —— **56**
3.2 Input-Output Multipliers —— **58**
3.2.1 The principal models —— **58**
3.2.2 A Model of Recovering the Sectorial Greenhouse Gas Emission Structure —— **61**
3.3 Updating and Forecasting I-O Tables —— **64**
3.3.1 Generalities —— **64**
3.3.2 RAS Formalism and its Limits —— **65**
3.3.3 Application: Updating an Aggregated EU I-O Matrix —— **68**
3.3.3.1 The RAS Approach —— **68**
3.3.3.2 The Entropy Approach —— **72**
3.3.3.3 I-O table Forecasting —— **73**
3.3.3.4 The Non-Extensive Cross-Entropy Approach and I-O Table Updating —— **75**
3.3.3.5 Forecasting an I-O Table —— **77**

3.3.4 Application: Forecasting the Aggregated 27 EU IO Coefficients —— **79**
3.3.5 Emission Coefficients Forecasting: A Theoretical Model —— **81**
3.4 Conclusions —— **83**
Bibliography – Part III —— **83**

PART IV: **Social Accounting Matrix**

1 Position of the Problem —— 88
2 A SAM as a Walrasian Equilibrium Framework —— 90
3 The Social Accounting Matrix (SAM) Framework —— 93
3.1 Generalities —— **93**
3.2 Description of a Standard SAM —— **94**
4 Balancing a SAM —— 97
4.1 Shannon-Kullback-Leibler Cross-Entropy —— **98**
4.2 Balancing a SAM through Tsallis-Kullback-Leibler Cross-Entropy —— **100**
4.3 A SAM as a Generalized Input-Output System —— **101**
4.3.1 A Generalized Linear Non-Extensive Entropy Econometric Model —— **102**
4.4 Input-Output Power Law (Pl) Structure —— **105**
4.5 Balancing a SAM of a Developing Country: the Case of the Republic of Gabon —— **107**
4.5.1 Balancing the SAM of Gabon by Tsallis Cross-Entropy Formalism —— **108**
4.6 About the Extended SAM —— **111**
5 A SAM and Multiplier Analysis: Economic Linkages and Multiplier Effects —— 117
5.1 What are the Economic Linkages and Multiplier Effects? —— **117**
5.1.1 A SAM Unconstrained Multiplier —— **119**
5.1.2 Equation System for Constrained SAM Multiplier —— **121**
5.1.3 On Modelling Multiplier Impact for an Ill-Behaved SAM —— **122**
Annex C. Proof of Economy Power Law Properties —— 122
Bibliography – Part IV —— **129**

PART V: **Computable General Equilibrium Models**

1 A Historical Perspective —— 134
2 The CGE Model Among Other Models —— 137
3 Optimal Behaviour and the General Equilibrium Model —— 138
3.1 Introduction —— **138**
3.2 Economic Efficiency Prerequisites for a Pareto Optimum —— **140**
4 From a SAM to a CGE Model: a Cobb-Douglas Economy —— 143
4.1 CGE Building Steps —— **143**
4.2 The Standard Model of Transaction Relationships —— **146**

5 Estimating the CGE Model through the Maximum Entropy Principle —— 151
5.1 Introduction —— **151**
5.2 Estimation Approach —— **153**
5.3 Application: Non-Extensive Entropy and Constant Elasticity of Substitution-Based Models —— **158**
5.3.1 Power Law and the Constant Elasticity of Substitution (CES) Function —— **159**
5.3.2 Parameter Outputs of the Tsallis Relative Entropy Model —— **163**
5.4 Conclusions —— **169**
Bibliography – Part V —— **169**

Part VI: **From Equilibrium to Real World Disequilibrium: An Environmental Model**

1 Introduction —— 176
2 Extending to an Environmental Model —— 179
2.1 The Model of Carbon Tax —— **181**
2.2 Carbon Tax Model and Double-Dividend Hypothesis —— **183**
3 Compensatory and Equivalent Variations: Two Types of Welfare Measurement —— 185
4 A Theoretical Example: CGE Model and Double-Dividend (DD)-Oriented Policies —— 187
5 Conclusions —— 190
Bibliography – Part VI —— **190**

Part VII: **Concluding Remarks**

Appendix —— 198
Annex C. Computational Aspects of Using GAMS —— 198
1.1 Introduction to GAMS —— **198**
1.2 Formulation of a Simple Linear Problem in GAMS —— **198**
1.3 Application to Maximum Entropy Models —— **199**
1.3.1 The Jaynes Unbalanced Dice Problem —— **199**
1.3.2 Non-Extensive Entropy Econometric Model —— **200**
Annex D. Recovery of Pollutant Emissions by Industrial Sector and Region: an Instructional Case —— 204
1.1 Introduction —— **204**
1.2 Recovery of Pollutant Emission by Industrial Sector and Region: the Role of the *Prior* —— **204**

Index of Subject —— 208
Index of Authors —— 209

Acknowledgements

My thanks are addressed to De Gruyter, the publisher, for its decisive financial support for publishing this monograph. I would like to acknowledge De Gruyter editors for their diligent and meticulous work to render this book scientifically understandable. I thank the University of Information Technology and Management in Rzeszów for its assistance and its support of my research. I am grateful to my former supervisor, M. Rocki, for his precious advice during the writing of the document. My particular thanks go to L. Tomasiewicz and J. Boratyński, both from the Łódź Department of Econometrics, for their technical advice on Polish National Accounts. Impressive work has been carried out by J. Kwapień and S. Drożdż, both specialists in thermodynamic statistics, who have identified omissions and errors over a large part of the book during the middle stage of its redaction. To them, I express all my gratitude. I express my thanks to A. Jakimowicz for his precious review comments. My thanks go also to the research institution Clersé at the University of Lille I in France, and in particular to N. Vaneecloo for having arranged my fruitful stay in that institution during my sabbatical year 2010-2011. I would like to thank F. Scheuren for having particularly increased my interest in the theory of information through the rich literature on the subject which was made available to me. Without the many months of sacrifice of M. Bloom for making this document readable, the publication of this work would not have been possible. Many thanks to all the other persons from my institution who at different levels have contributed directly or indirectly to improve the quality of this book. To all these people, I am most grateful, and for any errors or shortcomings which may remain in this book, they are absolved from all blame.

Summary

This book provides a new and robust power-law (PL)-based, non-extensive entropy econometrics approach to the economic modelling of ill-behaved inverse problems. Particular attention is paid to national account-based general equilibrium models known for their relative complexity.

In theoretical terms, the approach generalizes Gibbs-Shannon-Golan entropy models, which are useful for describing ergodic phenomena. In essence, this entropy econometrics approach constitutes a junction of two distinct concepts: Jayne's maximum entropy principle and the Bayesian generalized method of moments. Rival econometric techniques are not conceptually adapted to solving complex inverse problems or are seriously limited when it comes to practical implementation.

In recent years, PL-based Tsallis entropy has been applied in many fields. Its popularity can be attributed to its ability to more accurately describe heavy tail, non-ergodic phenomena. However, the link between PL and economic phenomena has been neglected—probably because the Gaussian family of laws are globally sufficient for time (or space) aggregated data and easy to use and interpret. Recent literature shows that the amplitude and frequency of macroeconomic fluctuations do not substantially diverge from many extreme events, natural or human-related, once explained at the same time or space-scale by PL. In particular, in the real world, socioeconomic rare events may, through long-range correlation processes, have higher impact than more frequent events could. Because of this and based on existing literature, this monograph proposes an econometric extension called *Non-extensive Entropy Econometrics* or, using a less technical expression, *Superstar-Generalised Econometrics*.

Recent developments in information-theoretic built upon Tsallis non-additive statistics are powerful enough to put established econometric theory in question and suggest new approaches. As will be discussed throughout this book, long-range correlation and observed time invariant scale structure of high frequency series may still be conserved—in some classes of non-linear models—through a process of time (or space) aggregation of statistical data. In such a case, the non-extensive entropy econometrics approach generally provides higher parameter estimator efficiency over existing competitive econometrics procedures. Next, when aggregated data converge to the Gaussian attractor, as generally happens, outputs from Gibbs-Shannon entropy coincide with those derived through Tsallis entropy. In general, when the model involved displays less complexity (with a well-behaved data matrix) and remains closer to Gaussian law, computed outputs by both entropy econometrics approaches should coincide or approximate those derived through most classical econometric approaches. Thus, the proposed non-ergodic approach could at least be as good as the existing estimation techniques. On empirical grounds, it helps in ensuring stability of the estimated parameters and in solving some classes of, up to now, intractable non-linear PL-related models. Furthermore, the approach remains one of the most appro-

priate for solving all classes of inverse problems, whether deterministic or dynamic. It is a more general approach. Finally, this approach helps us better assess—thanks to the Tsallis-q parameter—the interconnection level (complexity) between economic systems described by the model.

Consequently, this book aims at providing a new paradigm for econometric modelling through non-extensive (cross) entropy information-theoretic. Reaching this goal requires some intermediary results obtained through a synthesis of the existing, sometimes sparse literature. There are, then, methodological issues to address. Among these is the application of non-extensive entropy to low frequency time series. This constitutes a new challenge and must be clarified. Next, generalizing Gibbs-Kullback-Leibler information divergence to the Tsallis non-ergodic econometric model with different constraining moment formulations in both classes of entropy model will require special attention since we are not aware of any publications on the subject. Another important intermediary result of this work will be the proposition of a new theorem linking PL and macroeconomics on both the supply and demand sides. Its demonstration will provide new keys for carrying out further Tsallis entropy econometric modelling. Finally, we will provide an *ad hoc* statistical inference corresponding to the new modelling approach presented here.

The first part of the monograph presents basic targets and principal hypotheses.

In the second part, we present definitions and quantitative properties of statistical theory of information. Progressively, a link between the statistical theory of information and the generalized ill-posed inverse problem is established. After having shown the properties of the Shannon-Jaynes maximum entropy principle in detail, techniques for solving ill-behaved problems, from the Moore-Penrose generalized inverse problem to non-extensive entropy, are compared. Intrinsic relationships between both forms of Shannon-Jaynes[1] and Tsallis entropies are also shown. After having presented Kullback-Leibler information divergence, a generalization of this concept to non-extensive entropy is developed. A general linear non-extensive entropy econometric model is then introduced. It will play an important role for models to be developed in subsequent chapters. Next, an inferential formalism for parameter confidence interval area is proposed. This part is concluded with an applications example: the estimation of a Tsallis entropy econometrics model using the case of labour demand anticipation with a time series, error-correction model. Its outputs are compared with those of other approaches through Monte-Carlo simulations.

The third and fourth parts of the book—and, to a certain extent, the fifth part—are closely related to each other since a social accounting matrix can be seen as a kind of input-output transaction matrix generalization. The separation of these two

---

1 Here we prefer to shorten the name of this form of entropy. Scientists who have contributed to this form of entropy are many and cannot all be mentioned. It could be 'succinctly' named "Gibbs-Shannon- Jaynes-Kullback-Leibler Golan entropy econometrics."

parts has avoided highly horizontal and vertical subdivisions in the book, thereby preserving the clarity of the study. These two parts provide economic applications of statistical theory presented through Part II. Part III focuses on updating and forecasting national accounts tables. In particular, a new efficient approach to forecast input-output tables—or their extended forms—is set forth. The RAS approach is presented as a competing technique with empirical application and comments. To show one of the possible fields of entropy model implementation, we provide an ecological model to be solved as an inverse problem.

After having proposed a theorem linking PL distribution and the macroeconomic aggregative structure of national accounts, the problem of balancing a social accounting matrix (SAM) in the context of non-ergodicity is posed and solved in Part IV. The example presented deals with the actual problems of updating a SAM in real-world conditions.

In Part V, a computable general equilibrium (CGE) model is presented as a national account-related model. Two important concepts are discussed in the context of optimum property that both of them convey: the maximum entropy principle and the Pareto-optimum. Next, we open a short, epistemological discussion on two competitive and frequently confused estimation approaches, the Bayesian approach and the maximum entropy principal. An approach using non-extensive relative entropy for parameter estimation in the case of a constant elasticity of substitution (CES) function is proposed through the presentation of the CGE model.

To show the extensions of the standard national accounts table and to go beyond the general equilibrium framework, an environmentally extended social accounting matrix and a subsequent theoretical model displaying externalities are presented in Part VI. Finally, a carbon tax and double dividend theory model is presented and its social welfare impact is derived as well.

The last part of the book concludes with the principal findings and proposes areas for further investigation and research.

Two examples are provided in Annex C and D. The first concerns the use of GAMS as a platform for economic programming. The second presents some hints for solving inverse problems in the context of the proposed model.

To enable readers to better understand the results in the different chapters, they are accompanied by detailed examples or case studies and summarizing comments. As such, this book can be an ideal reference for students and researchers in many disciplines (infometrics, econometrics, statistics, national accounting, optimal control, etc.) interested in becoming familiar with approaches that reflect the most recent developments in statistical theory of information and their application for stochastic inverse problem modelling. Last but not least, the discussion in this book is limited to technical issues; it does not cover the philosophical implications of non-extensive entropy, whether general or within the discipline of economics.

PART I: **Generalities and Scope of the Book**

1 Generalities

1.1 Information-Theoretic Maximum Entropy Principle and Inverse Problem

1.1.1 Information-Theoretic Maximum Entropy Principle

According to recent literature (Golan, Judge, & Miller, 1996; Golan, 2008), the information-theoretic maximum entropy principle is a coincident junction of two lines of research: inferential statistics and statistical thermodynamics.

The first line of research emerged in the beginning of the 18th century through the work of Bernoulli (Jaynes, 1957; Halmos & Savage, 1949; Bayes, 1763; and Laplace, 1774). They developed the Principle of Sufficient Reason, which consists of determining the state of the system on the basis of limited information (moments) from a subsystem. This principle was later extended in the last century by Jeffreys (1946), Cox (1946), and Jaynes (1957b) to the principle of "not telling more than you know," thus suggesting the necessity of avoiding additional hypotheses imposed merely to simplify the problem to be solved. The purpose of all of the above authors' research was to retrieve characteristics of a general population on the basis of limited information from a possibly non-representative sample of that population, out of risky or non-convenient hypotheses.

The second line of research is represented, amongst others, by Maxwell (1871), Boltzmann (1871), Cauchy (1855), Weierstrass (1886), Lévy and Gibbs (Gibbs, 1902), Shannon (1948), Jaynes (1957, 1957b), Rényi (1961), Bregman (1967), Mandelbrot (1967), Tsallis (1988). Its main objective was to provide mathematical formalism to statistical modelling of physical information related to natural phenomena. Thanks to the celebrated work of Tsallis (1988), on non-extensive thermodynamics[2], this second line elegantly extended its multidisciplinary applications to "auto-organized systems" and to the social sciences, particularly in financial fields.

The ascent and development of the post-war information theory-based, maximum entropy proposed by Shannon (1948) can be viewed as a major step toward the rapid extension of the discipline. Less than a decade was needed to develop the information-theoretic principles of statistical inference, inverse problem solution methodology based on Gibbs-Shannon maximum entropy, and its generalizations by Kullback and Leibler (1951), Kullback (1959) and Jaynes (1957b). The above authors developed, in particular, fundamental notions in statistics, such as sufficiency and efficiency

2 Currently, this theory—undoubtedly the best—generalizes Boltzmann-Gibbs statistics for describing the case of anomalous systems characterized by non-ergodicity or metastable states. It thus better fits dynamic correlation of complex systems and can be better explained (e.g., Douglas, 2006), amongst many others.

(Halmos & Savage, 1979), a generalization of Cramer-Rao inequality (e.g., Kullback, 1959) and the introduction of a general linear model as a consistency restriction (Heckelei et al., 2008) through Bayesian philosophy. Thus, it became possible to unify heterogeneous statistical procedures via the concepts of information theory. Lindley (2008), on the other hand, had provided the interpretation that a statistical sample could be viewed as a noisy channel (Shannon's terminology) that conveys a message about a parameter (or a set of parameters) with a certain prior distribution. This new interpretation extended application of Shannon's ideas to statistical theory by referring to the information in a statistical sample rather than in a message.

Over the last two decades the literature concerned with applying entropy in social science has grown considerably and disserves closer attention. On one side, Shannon-Jaynes-Kullback-Leibler-based approaches are currently used for modelling economic phenomena competitively with classical econometrics. A new paradigm in econometrical modelling is taking place and finds its roots in the influential work of Golan, Judge, and Miller (1996). The present monograph constitutes an illustration of this.

As mentioned above, this approach is particularly useful in the case of solving inverse problems or ill-behaved matrices when we try to estimate parameters of an econometric model on the basis of insufficient information from an observed sample, and this estimation may concern the behaviour of an individual element within the system.

Insufficient information implies that we are trying to solve an ill-posed problem, which plausibly can arise in the following cases:

— data from sampling design are not sufficient and/or complete due to technical or financial limitations—*small area official statistics* could illustrate this situation;

— non-stationary or non-co-integrating variables are resulting from bad model specification;

— data from the statistical sample are linearly dependent or collinear for various reasons;

— Gaussian properties of random disturbance are put into question due to, amongst many others things[3], systematic errors from the survey process;

— the model is not linear and *approximate* linearization remains the last possibility;

— aggregated (in time or space) data observations hide a very complex system represented, for instance, by a PL distribution, and multi-fractal properties of the system may exist.

3 It is not excluded that distribution law may be erroneously applied since, for instance, randomness is dependent on the experimental setup or the sophistication of the apparatus involved in measuring the phenomenon (Smith, 2001).

Using the traditional econometrical approaches in one or more of the above cases—without additional simplifying hypotheses—could lead to various estimation problems owing to the nonexistence of a bounded solution or the instability of estimates. Consequently, outputs from traditional econometrical approaches will display, at best, poor informative parameters. In the literature, there are other well-known techniques to cope with inverse problems or ill-conditioned data. Among them, two popular techniques deserve our attention: the bi-proportional RAS approach (and its variants), particularly used for updating or forecasting input/output matrices (Parikh, 1979) and the Moore-Penrose pseudo-inverse technique, useful for inverting irregular matrices (e.g., Green, 2003, p. 833). In spite of their popularity, both techniques present serious drawbacks in empirical investigations. In fact, the RAS techniques, in spite of their divergence information nature, remain less adapted to solving stochastic problems or to optimizing the information criterion function under a larger number of different prior constraining data. Since Moore-Penrose generalized inverse ensures a minimum distance (Y-BX) only when the matrix B has full rank, it will not reflect an optimal solution in other cases. Golan et al. (1996) have clearly shown higher efficiency of Shannon maximum entropy econometrics over the above cited methods in recovering unknown information when data or model design is poorly conditioned. The suggested superiority stands on the fact that it combines and generalizes maximum entropy philosophy (as in the second law of thermodynamics) and statistical theory of information attributes as a Bayesian information processing rule. As demonstrated convincingly by Golan (1996, 2006), Shannon entropy econometrics formalism may generalize least squares (LS) and the maximum likelihood (ML) approaches and belongs to the class of Bayesian method of moments (BMOM). It is worthwhile to point out that in the coming chapters many cases of cross-entropy (or minimum entropy) formalism will be used in place of maximum entropy. This is because, in this study, many problems to be treated involve information measuring in the context of the Kullback-Leibler framework.

This monograph does not intend to treat the case of high frequency series for which a rich literature already exists. We invite readers interested in the case of high frequency series to see, for instance, J.W. Kantelhardt (2008) for testing for the existence of fractal or multi-fractal properties, suggesting the case of a PL distribution.

1.2 Motivation of the Work

1.2.1 Frequent Limitations of Shannon-Gibbs Maximum Entropy Econometrics

In spite of a growing interest in the research community, some incisive critics have come forward to address Shannon-based entropy econometrics (e.g., Heckelei et al., 2008). According to some authors, generalized maximum entropy (GME) or cross-entropy (GCE) econometrical techniques face at least three difficulties. The first is related to the specification and interpretation of prior information, imposed via the use of discrete support points, and assigning prior probabilities to them. The authors argue that there are complications that result from the combination of priors and their interaction with the criterion of maximum entropy or minimum cross-entropy in determining the final estimated *a posteriori* probabilities on the support space. The second group of criticisms questions the sense of the entropy objective function once combined with the prior and data information. The last problem, according to the same authors, refers to computational difficulties owing to the mathematical complexity of the model with an unnecessarily large number of parameters or variables.

Concerning the first criticism, the problem—selecting a prior support space and prior probabilities on it—exists since estimation outputs seem to be extremely sensitive to initial conditions. However, when there is a theory or some knowledge about the space on which parameters are supposed to be staying, the problem becomes tractable. In particular, when we have to estimate parameters in the form of ratios, the performance of entropy formalism is high. To this counterargument, it is worthwhile to add that GME or GCE formalism constitutes an approach based on the Bayesian efficient processing rule and, as such, prior values are not fixed constraints of the model; they combine and adapt with respect to other sets of information (e.g., consistency function) added to the model to update a new parameter level in the entropy criterion function.

The second problem concerns questioning the sense or interpretability of output probabilities from the maximum entropy criterion function once combined with real world probability-related restrictions. One cannot comment on this problem without making reference to the important contribution of Jaynes (1957, 1957b), who proposed a way to estimate unknown probabilities of a discrete system in the presence of less data point observations than parameters to be estimated through the celebrated example of Jaynes dice. Given a set of all possible ways of distribution resulting from all micro-elements of a system, Jaynes proposed using the one that generates the most "uncertain"[4] distribution. To understand this problem, the question becomes a matter of combining philosophical interpretation of the maximum entropy principle with that of Jaynes' formulation in the context of Shannon entropy. Depending on the type of entropy[5] considered, output estimates will have slightly different meaning. However, all interpretations refer to parameter values that assure a long-run, steady-

4 Here we are in the realm of the second law of thermodynamics, which stipulates, in terms of entropy, that natural equilibrium of any set of events is reached once disorder inside them becomes optimal. This results from their property of having equal (ergodic system) odds to occur. In that state, we reach the maximum uncertainty about which event should occur in the next trial.

5 Later, for comparison, properties of the most well-known types of entropy in the literature will be presented.

state equilibrium of the system (relations defined by the model) with respect to data and other knowledge at hand, usually in the form of moments and/or normalization conditions. Owing to maximum entropy alone, the more consistent moments are or the more other *a priori* information binds, the more output probabilities will differ from those in a uniform distribution. Considering the above, interpretation of the maximum entropy model is far removed from interpretation of the classical model, especially in the case of the econometric linear model where estimates mean a change in the endogenous variable due to unitary change in an explicative variable, that is, in *ceteris paribus* conditions.

The last criticisms concern the burden arising from the computational and numerical process—a problem common to all complex, nonlinear systems. Thanks to recent developments of computer software, this problem is now less important.

In many empirical studies that attempt to solve inverse problems, the Shannon entropy-based approach is relatively efficient in recovering information. However, gaining in parameter precision requires good design of the prior. In particular, the point support space must fit into the space of the true population parameter values. As Golan et al. (1996) have shown, when prior design is weak, outputs of Shannon entropy econometrics will produce approximately the same parameter precision as traditional econometrical methods, such as LS or the ML, which means Shannon entropy could discount information not fitting the maximum entropy principle as expected.

The above criticisms of the Shannon entropy econometrics model remain relatively weak as has been shown through the preceding discussion.

According to us, the main drawback related to that form of model is due to the analytical function of constraining moments. In fact, as already suggested, long-range correlation and observed time invariant scale structure of high frequency series may still be conserved—in some classes of non-linear models—through a time—or space—aggregation process of statistical data. This raises the question of why this study proposes a new approach of Tsallis non-extensive entropy econometrics.

The next section provides a first answer by showing potential theoretical and then empirical drawbacks of the Shannon-Gibbs entropy model and potential advantages from the PL-related Tsallis non-extensive entropy approach.

1.2.2 Rationale of PL-Related Tsallis Entropy Econometrics and Low Frequency Series

This section presents the essence of the scientific contribution of this monograph to econometric modelling. For a few decades, PL has confirmed its central role in describing a large array of systems, natural and manmade. While most scientific fields have integrated this new element into their analytical approaches, econometrics and hence, economics globally, is still dwelling—probably for practical reasons—

on the Gaussian fundamentals. This study takes a step forward by introducing Tsallis non-extensive entropy to low frequency series econometric modelling. The potential advantages of this new approach will be presented, in particular, its capacity to analytically solve complex PL-related functions. Since any mathematical function form can be represented by a PL formulation, the importance of the proposed approach becomes clear. To be concrete, one of the complex nonlinear models is the fractionally integrated moving average (ARFIMA) model, which, to our knowledge, has remained non-tractable using traditional statistical instruments. An empirical application to solve such a class of models will be implemented at the end of Part V of this book.

According to several studies (Bottazzi et al., 2007), (Champernowne, 1953), (Gabaix, 2008), a large array of economic laws take the form of a PL, in particular macroeconomic scaling laws, distribution of income, wealth, size of cities and firms[6], and distribution of financial variables such as returns and trading volume. Ormerod and Mounfield (2012) underscore a PL distribution of business cycle duration. Stanley et al. (1998) have studied the dynamics of a general system composed of interacting units, each with a complex internal structure comprising many subunits, where the subunits grow in a multiplicative way over a period of twenty years. They found that this system followed a PL distribution. It is worthwhile to note the similarity of such a system with the internal mechanism of national accounts tables, such as a SAM, also composed of interacting economic sectors, each with a complex internal structure defined by firms exercising similar business. Ikeda and Souma (2008) have made an international comparison of labour productivity distribution for manufacturing and non-manufacturing firms. A PL distribution in terms of firms and sector productivity was found in US and Japanese data. Testing the Gibrat's law of proportionate effect, Fujiwara et al. (2004) have found, among others things, that the upper-tail of the distribution of firm size can be fitted with a PL (Pareto-Zipf law). The list of PL evidence here is limited to social science.

Since this study focuses on the immense potentiality of PL-related economic models, PL ubiquity in the social sciences will be underscored and a theorem showing the PL character of national accounts in its aggregate form will be presented.

In line with the rationale for the proposed methodology detailed below, the following from recent literature is evidence of entropy:

– Non-extensive entropy, as such, models the non-ergodic systems which compound Levy[7] instable phenomena[8] converging in the long range to the Gaussian basin of attraction. In the limiting case, non-extensive entropy converges to Shannon Gibbs entropy.

---

6 See (Bottazzi et al., 2007) for different standpoints on the subject.
7 Shlesinger, Zaslavsky, & Klafter, Strange Kinetics, 1993.
8 Shlesinger et al., Lévy Flights and Related Topics in Physics, 1995.

- PL-related Tsallis entropy should remain, even in the case of a low frequency series, a precious device for econometric modelling since the outputs provided by the exponential family law (e.g., the Gibbs-Shannon entropy approach) correspond to the Tsallis entropy limiting case when the Tsallis-q parameter equals unity.
- A number of complex phenomena involve long-range correlations which can be seen particularly when data are time scale-aggregated (Drożdż & Kwapień, 2012), (Rak et al., 2007). This is probably because of the interaction between the functional relationships describing the involved phenomena and the inheritance properties of a PL or because of their nonlinearity. Delimiting the threshold values for a PL transition towards the Gaussian structure (or to the exponential family law) as a function of the data frequency amplitude is difficult since each phenomenon may display its own rate of convergence—if any—towards the central theorem limit attractor.
- Systematic errors from statistical data collecting and processing may generate a kind of tail queue distribution. Thus, a systematic application of the Shannon-Gibbs entropy approach in the above cases—even on the basis of annual data—could be misleading. In the best case, it can lead to unstable solutions.
- On the other hand, since non-extensive Tsallis entropy generalizes the exponential family law (Nielsen & Nock, 2012), the Tsallis-q entropy methodology fits well with high or low frequency series.

In the class of a few types of entropy displaying higher-order entropy estimators able to generalize the Gaussian law, Tsallis non-extensive entropy has the valuable quality of concavity–and then stability—along the existence interval characterizing most real world phenomena. As far as the q-generalization of the Kullback-Leibler (K-L) relative entropy index is concerned, it conserves the same basic properties as the standard K-L entropy and can be used for the same purpose (Tsallis, 2009).

The above-enumerated points imply that in cases where the assumed Levy law complexity is not verified by empirical observation, outputs from the non-extensive entropy model converge with those derived from Shannon entropy. In other words, errors which involve taking a sample as if it were PL-driven has no consequence on outputs if the truth model belongs to the Gaussian basin of attraction. This explains why in most empirical applications—but by no means all—both forms of entropy provide similar results and the entropic Tsallis-q complexity parameter then tends to converge to unity, revealing the case of a normal distribution. Empirical examples will be presented at the end of this document, and the strength of Tsallis maximum entropy econometrics will be demonstrated in different contexts.

In summary, the following are entropy function regularities:

- The Tsallis entropy model generalizes the Shannon-Gibbs model, which constitutes a converging case of the former for the Tsallis-q parameter equal unity.

– The Shannon-Gibbs model fits natural or social phenomena displaying Gaussian properties.
– PL high frequency time (space) series scaling—aggregating—does not always lead to Gaussian low frequency time (space) series. Additionally, the rate of convergence from the PL to the Gaussian model, if any, varies according to the form of the function used.

Is it judicious to replace Shannon-Gibbs entropy modelling by Tsallis non-extensive entropy for empirical applications?

The answer is yes, and this is the motivation for this study. There are at least three expected advantages to introducing Tsallis non-extensive econometric modelling:

1. A data generating system characterized by a low—or no—convergence rate from PL to Gaussian distribution only becomes analytically tractable when using Tsallis entropy formalism. (This will be proven through an econometrical model with constant substitution elasticity and then considered as an inverse problem to be estimated later.)
2. The Tsallis entropy model displays higher stability than the Shannon-Gibbs, particularly when systematic errors affect statistical data.
3. The Tsallis-q parameter presents an expected advantage of monitoring complexity of systems by measuring how far a given random phenomenon is from the Gaussian benchmark. In addition to other advantages, this can help draw attention to the quality of collected data or the distribution involved.

The choice of national accounts-related models for testing the new approach of non-extensive entropy econometrics is motivated by the empirical inability of national systems of economic information to provide consistent data according to macroeconomic general equilibrium. As a result, national account tables are generally not balanced unless additional—often contradictory—assumptions are applied to balance them. However, following the principle of not adding (to a hypothetical truth) more than we know, it remains preferable to deal with an *unbalanced* national accounts table. Trying to balance such a table implies that we are faced with ill-behaved inverse problems. According to the existing literature, and as will be seen through this monograph, entropy formalism remains the best approach to solving such a category of complex problems. The superiority of Tsallis non-extensive entropy econometrics over other known econometrical or statistical procedures results from its capacity to generalize a large category of most known laws, including Gaussian distribution.

1.3 National Accounts-Related Models and the Scope of this Work

Under the high frequency series hypothesis, we postulate that social and economic activities are characterized by complex behavioural interactions between socio-economic agents and/or economic sectors. Recent, *Big Data* for Official Statistics may illustrate such a complexity. This could mean that the supposed extreme events may appear systematically more (or less) frequently than expected (Gaussian scheme), implying internal and aggregated long-range correlation (over time, space, or both). The maximum entropy principle is best suited to estimating ill-behaved inverse problems and, in particular, models with ratios or elasticity as parameters. In this latter case, as we will see later, the support space area for unknown parameters coincides with the probability area over the space from zero to unity. Fortunately enough, due to its macroeconomic consistency, national account table structure reflects this property. In empirical macroeconomic investigations, the national accounts system of information plays a crucial role for modelling as it guarantees internal coherence of macroeconomic relations. Numerical information is embodied inside comprehensive statistical tables or balance sheets displaying algebraic properties of a matrix. Having in mind an economic or statistical inference investigation, mathematical treatment of information compounded inside these matrices is carried out by economists or statisticians on the basis of *a priori* information at hand. When such matrices are algebraically regular, traditional inverse methods can be applied to solve the problem of, for instance, estimating parameters that define relationships between the endogenous variable and its covariates. Nevertheless, in the social sciences, causality relationships linking both variables seldom have a one-to-one correspondence. In many cases, two or more different inputs or causes can lead to the same output or effect. Such different causal concomitances for the same output render the social or economic model indeterminate. In such cases, the recovery of a data generating system from the observed finite sample becomes impossible using the traditional statistical or econometric devices, such as the standard maximum likelihood method or the generalized method of moments. On mathematical grounds, this may result from an insufficient number of model data points with respect to the number of parameters to estimate. Such a sample is said to be ill-behaved. This situation leads to the lack of an optimal solution sought. Collinear variables, inadequate size of a small sample, or the poor quality of statistical data may lead to the same difficulties. Finally, taking into account the above deficiencies and anomalies, modellers have to deal with ill-behaved inverse problems most of the time. Following what has been said above, this monograph targets developing a robust approach generalizing Kullback-Leibler-Shannon entropy for solving inverse problems related to national account models in a way that reflects the complex relationships between economic institutions and/or agents. Statistical data from such complex interrelations are usually difficult to collect, incomplete, and defective. Additionally—and this may be one of the most important points—modelling national account table-related information involves some class of nonlinear functions, otherwise only solvable using the PL model; thus, non-ergodic situations are involved.

The next area of national accounts modelling to be treated in this monograph is:

- Updating an input/output table when the problem is posed as inverse, with the possibility of adding extra sample information to the model in the form of an *a priori* and without any additional assumption;
- Forecasting an input/output table or its extended forms, such as the social accounting matrix (SAM), solely on the basis of yearly published national accounts concerning sectorial elements of final demand and gross domestic product;
- Deriving backward or forward multiplier coefficient impact on the basis of insufficient pieces of information;
- Demonstrating a method to forecast a sectorial energy final demand and total pollutants emission by production, the basis of an environmentally extended input/output table when basic information is missing;
- Presenting a computable general equilibrium model using the maximum entropy approach instead of calibration techniques to derive the parameters of CES functions;
- Estimating other nonlinear economic functions as inverse problems and conducting Monte Carlo experiments to test Tsallis entropy econometrics outputs;
- Presenting in detail, across different chapters, national account-related general equilibrium models before coming back to inverse problem solution techniques as suggested above.

The reader should be enriched not only by techniques for solving complex inverse problems but also by a thorough examination of different aspects of national account updating and modelling in the Walrasian spirit. To render the models presented here more consistent, emergent elements on an environmentally extended system of accounts will be included along with their impact on the general equilibrium framework and the optimum Pareto or social welfare.

Bibliography – Part I

Bayes, T. (1763). An essay towards solving a Problem in the Doctrine of Chances, Phil. Trans. 53, 370–418 (1763). http://rstl.royalsocietypublishing.org/content/53/370

Bottazzi, G. et al. (2007). Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics* 29, pp. 137–159.

Bregman, L.M. (1967). The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *Computational Mathematics and Mathematical Physics* 7(3), pp. 200–217.

Champernowne, D.G. (1953). A Model of Income Distribution. *The Economic Journal* 63(250), pp. 318–351.

Cox, R.T. (1946). Probability, Frequency, and Reasonable Expectation. *Am. Jour. Phys.,* 14, pp. 1–13.

Douglas P. et al. (2006). Tunable Tsallis Distributions in Dissipative Optical Lattices. *Phys. Rev. Lett.* 96, 110601.

Drożdż, S., & Kwapień, J. (2012). Physical approach to complex systems. *Physica reports*, pp. 115–226.

Fujiwara, Y. et al. (2004). Gibrat and Pareto–Zipf revisited with European firms. *Physica* A 344*, 1–2*, pp. 112–116.

Gabaix, X. (2008, September). *Power Laws in Economics and Finance*. Retrieved from NBER: http://www.nber.org/papers/w14299

Gibbs, J.W. (1902). *Elementary principles in statistical mechanics.* New York, USA: C. Scribner's Sons incl.

Golan, A. (2006), Information and Entropy Econometrics — A Review and Synthesis, *Foundations and Trends in Econometrics*, vol 2, no 1–2, pp 1–145.

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

Green, W. (2003). *Basics of Econometrics, 5th edition.* NY: Prentice Hall.

Halmos, P.R., & Savage, L.J. (1949). Application of the Radon-Nikodyma theorem to the theory of sufficient statistics. *Annals of Math. Stat. 20*, pp. 225–241.

Hartley, R.L. (1928, July). Transmission of Information. *Bell System Technical Journal*.

Heckelei, T. et al. (2008). *Bayesian alternative to generalized cross entropy solutions for underdetermined econometric models.* Bonn: 2008/2, University of Bonn.

Ikeda, Y., & Souma, W. (2009). *International comparison of Labor Productivity Distribution for Manufacturing and Non-manufacturing Firms*. Progress of Theoretical Physics Supplement, 179, pp. 93, Oxford University Press.

Jaynes, E.T. (1957). Information Theory and Statistical Mechanics. *Physical Review*, pp. 620–630.

Jaynes, E.T. (1957b). *Probability Theory: The Logic Of Science.* USA: Washington University.

Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems. (R.S. London, Ed.) *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 186 (1007), pp. 453–461.

Kantelhardt, J.W. (2008). *Fractal and multifractal time series.* Wisconsin, USA: Institute of Physics, Martin Luter University.

Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* 22, pp. 79–86.

Kullback, S. (1959). *Information theory and statistics.* NY: John Wiley and Sons.

Laplace, S.P. (1774). Memoir on the probability of causes of events. (English translation by S.M. Stigler, Ed.) *Statist. Sci. 1(19), Mémoires de Mathématique et de Physique, Tome Sixième*, 364–378.

Lindley, D. (2008). *Uncertainty: Einstein, Heisenberg, Bohr, and the Struggle for the Soul of Science.* NY: Anchor.

Mandelbrot, B. (1967). How Long Is the Coast of Britain? Statistical Self-Similarity and Fractional Dimension. *Science,* 156(3775), pp. 636–638.

Maxwell, J.C. (1871). *Theory of Heat.* New York: Dover.

Nielsen, F., & Nock, R. (2012). A closed-form expression for the Sharma-Mittal entropy of exponential families. *Journal of Physics A: Mathematical and Theoretical 45*, p. 3.

Ormerod, P., & Mounfield, C. (2012). Power law distribution of the duration and magnitude of recessions in capitalist economies: Breakdown of scaling. *Physica* A293, pp. 573–582.

Parikh, A. (1979). Forecasts of Input-Output Matrices Using the R.A.S. Method. *The Review of Economics and Statistics 61, 3*, pp. 477–481.

Rak, R. et al. (2007). Nonextensive statistical features of the Polish stock market fluctuations. *Physica A 374*, pp. 315–324.

Rényi, A. (1961). On Measures of Entropy and Information. In Univ. of California (Ed.), *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability,* 1 (pp. 547–561). Berkeley, Calif.: University of California Press.

Kullback, S. (1959). *Information theory and statistics,* NY: John Wiley and Sons.

Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal,* 27, pp. 379–423 & 623–656.

Shlesinger, M. et al. (1995). *Lévy Flights and Related Topics in Physics.* Berlin: Springer.

Shlesinger, M., Zaslavsky, & Klafter. (1993). Strange Kinetics. *Nature* 363, p. 31–37.

Smith, J.D. (2001). Some Observations on the Concepts of Information-Theoretic Entropy and Randomness. *Entropy* 3[1], pp. 1–11.

Stanley, H. et al. (1998). Power Law Scaling for a System of Interacting Units with Complex Internal Structure. *Phys Rev Lett* 80, p. 7.

Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *Journal of Statistical Physics* 52, pp. 479–487.

Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World.* Berlin: Springer.

Weierstrass, K. (1886). *Abhandlungen aus der Functionenlehre [Treatises from the Theory of Functions].*Berlin, Germany: Julius Springer.

PART II: **Statistical Theory of Information and Generalised Inverse Problem**

1 Information and its Main Quantitative Properties

1.1 Definition and Generalities

This chapter constitutes the base of the next theoretical formalism to be developed in this part of the book. The connection between the Bayesian rule and Kullback information divergence is first envisaged. This will permit a better understanding of Shannon-Jaynes-Kullback cross-entropy. The next section will deal with the connection between Shannon-Gibbs entropy and non-extensive (Tsallis) entropy. Finally, the generalized non-extensive cross-entropy will be presented for further applications in the remaining parts of the work.

Many forms and measures of information exist. As far as parameters linked to data observations are concerned, one well-known measure of information was provided by R.A. Fisher in 1929. As will be clear below, the next can be log(n), explaining the sum of n hypotheses H_i, all uniformly distributed and known as Hartley's information measure (Hartley, 1928). Information theory has its mathematical roots in the concept of disorder or entropy in statistical mechanics. Kullback (1959) provides an extensive literature on the form and mathematics linking entropy and information theory. As mentioned, the next formal definition will be followed by theoretical and empirical extensions arising from the entropy principle.

Let us now develop a workable measure of information obtained through observation of an event having probability p. Our first problem is to ignore any particular features of the event and focus only on whether or not it happened. Thus we will think of an event as the observance of a symbol whose probability of occurring is p. Thus, the information will be defined in terms of the probability p.

Let us consider the probability spaces (χ, ϑ, μ_i), i = 1,2 as a basic set of elements $x \epsilon \chi$ (sample space) and the $\sigma -$ algebra ϑ, a collection of all possible sets of events from χ with the probability measure μ_i. Under general assumptions of the above probability measures, in particular those stating their absolute continuity with respect to one another, let $\lambda \equiv \mu_i$. By the Radon-Nikodym theorem (e.g., Loeve, 1955), there exist functions $f_i(x)$, $i = 1,2$, called generalized probability densities, $0 < f_i(x) < \infty$ $[\lambda]$ such that:

$$\mu_i(\mathrm{E}) = \int_{\mathrm{E}} f_i(x) d\lambda(x), \;\; i = 1,2, \tag{2.1}$$

for all E belonging to the $\sigma -$ algebra ϑ. Following Kullback (1959) and Halmos & Savage (1949), the symbol $[\lambda]$, pronounced "modulo λ", means that the assertion is true along with all the support space of events E except the case for E $\epsilon\, \vartheta$ and λ (E)=0.

In (2.1), the function $f_i(x)$ is also referred to as the Radon-Nikodym derivative. If the probability measure μ is absolutely continuous with respect to the probability measure λ and the probability measure ν is absolutely continuous with respect to μ, then the probability measure v is also absolutely continuous with respect to λ, and the Radon-Nikodym derivatives satisfy:

$$\frac{dv}{d\lambda} = \frac{dv}{d\mu} \cdot \frac{d\mu}{d\lambda}[\lambda]$$

The defined symbols above allow us to better derive the conceptual definition of information below as it will be understood in the coming chapters of this book.

Next, let H_i, $i = 1,2$, be the hypothesis that a variable is X from the statistical population with probability measure μ_i. Then, by applying Bayes's theorem, it follows that:

$$P(H_i \mid x) = \frac{P(H_i) f_i(x)}{P(H_1) f_1(x) + P(H_2) f_2(x)}[\lambda], \; i = 1,2, \tag{2.2}$$

After transformations with respect to logarithms of relative function densities $f_i(x)$, we obtain:

$$\log \frac{f_1(x)}{f_2(x)} = \log \frac{P(H_1 \mid x)}{P(H_2 \mid x)} - \log \frac{P(H_1)}{P(H_2)}[\lambda], \tag{2.3}$$

where: x is an element of X; $P(H_i)$ is the prior probability of H_i and $P(H_i \mid x)$ is the posterior probability of H_i. The logarithm in (2.3) stands for an information measure base unit (Hartley, 1928). The right-hand side of (2.3) is an informative measure resulting from the difference (positive or negative) between the logarithm of the odds in favour of H_i once observation of x has occurred and before it occurred.

Thus, following Kullback, one defines the logarithm of the likelihood ratio,

$$\log \frac{f_1(x)}{f_2(x)},$$

as the information in $X = x$ for discrimination in favour of H_1 against H_2. An interesting alternative definition of information after (2.3) is the weight of evidence for H_1 given x (Kullback, 1959), (Good, 1963). Next, most informative is the mean information for discrimination in favour of H_1 against H_2 given $x \in E \in \vartheta$, for μ_1, which is defined as follows:

$$I(\mu_1 : \mu_2) = \int \log \frac{f_1(x)}{f_2(x)} d\mu_1(x) = \int f_1(x) \log \frac{f_1(x)}{f_2(x)} d\lambda(x) =$$

$$= \int \log \frac{P(H_1 \mid x)}{P(H_2 \mid x)} d\mu_1(x) - \log \frac{P(H_1)}{P(H_2)} \tag{2.4}$$

with $= d\mu_i(x) = f_1(x) d\lambda(x)$.

Here one has treated the general case when E represents the entire sample space χ and then must not appear as support space for integration (see 2.1). The last member in (2.4) is the difference between the mean value, with respect to μ_1, of the logarithms of the posterior and prior odds of the hypotheses. Following Savage (1954), Kullback (1959), *I(1:2)* could be referred to as the information of μ_1 with respect to μ_2.

Let us extend the above general definition of information to some known cases. Suppose we have a set (categories) of hypotheses, $H_i = 1, 2,.., n$ and that from observation, we can infer with certainty which hypothesis is true. Then the mean information in an observation about H is the mean value of -log $P(H_i)$, that is,

$$P(H_1) \log P(H_1) - P(H_2) \log P(H_2) - \ldots - P(H_n) \log P(H_n). \tag{2.5}$$

The expression in (2.5) above is called entropy of the H_i*'s* (e.g., Khinchin, 1957; Shannon, 1948). When hypotheses H_i are uniformly distributed (then equally probable) so that

P(Hi) = 1/n, i = 1…n, this leads to $-\sum_{i=1}^{n} P(H_i) \log P(H_i) = \log n,$

which turns out to be Hartley's information measure.

As shown below, an interesting applicability of (2.4) may concern the analysis of hypotheses H_i, $i = 1,2$, on dependency between variables x and y or on the measure of divergence between given hypotheses H_i. Presenting relationships between information discriminating measure and dependency between variables will be useful when we introduce an inferential approach for entropy econometrics models. In particular, measure of divergence constitutes, once again, the cornerstone of the present work in which *a priori* and *a posteriori* hypotheses will be recalled in many applicable analyses.

Suppose we have the entire sample space χ being the Euclidean space of two dimensions R^2 with elements $X = (x, y)$. Let us consider that under H_1 variables x and y are dependent with probability density $f(x, y)$ and that, under the alternative hypothesis H_2, both variables are independent with probabilities $g(x)$ and $h(y)$. In this case, we rewrite (2.4) as follows:

$$I(\mu_1 : \mu_2) = \iint f(x, y) \log \frac{f(x, y)}{g(x)h(y)} dxdy \tag{2.6}$$

Information measure $I(\mu_1 : \mu_2)$ is nonnegative (Kullback, 1959) and equal to zero if and only if $f(x, y) = g(x)\, h(y)$ [λ]. As such, it constitutes an informative indicator on dependency degree between x and y. Note that in the case of a bivariate normal density

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{\frac{1}{2}}}\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma^2{}_x}-2\rho\frac{xy}{\sigma_x\sigma_y}+\frac{y^2}{\sigma^2{}_y}\right)\right]$$

where hypothesis H_2 then represents the product of the normal densities as explained in (2.6), and finally one obtains:

$$I(\mu_1:\mu_2) = -\frac{1}{2}\log(1-\rho^2), \tag{2.7}$$

which indicates that in the case of bivariate normal distribution, as expected, the mean information is discriminatory in favour of H_1 (dependence) against H_2 (independence); that is, $I(\mu_1:\mu_2)$ is a function of the correlation coefficient ρ alone.

Following Jeffreys (1946) and Kullback (1959), if we define $I(2:1)$ as

$$I(\mu_1:\mu_2) = f_2(x)\log\frac{f_2(x)}{f_1(x)}d\lambda(x) \tag{2.8}$$

that is, the mean information from μ_2 for discrimination in favour of H_2 against H_1, one can define the divergence between hypotheses (noted ∇) by:

$$\nabla(H_1, H_2) = I(\mu_1:\mu_2) + I(\mu_2:\mu_1) = \int(f1(x)-f_2(x))\log\frac{f_1(x)}{f_2(x)}d\lambda(x)=$$

$$=\int\log\frac{\mathrm{P}(H_1\,|\,x)}{\mathrm{P}(H_2\,|\,x)}d\mu_1(x)-\int\log\frac{\mathrm{P}(H_1\,|\,x)}{\mathrm{P}(H_2\,|\,x)}d\mu_2(x) \tag{2.9}$$

Thus, $\nabla(H_1, H_2)$ measures the divergence between H_1 and H_2 or between μ_1 and μ_2. As such, it constitutes a measure of the difficulty of discriminating between them.

1.2 Main Quantitative Properties of Statistical Information

The approach undertaken here is axiomatic (Carter, 2011). It is worthwhile to note that we can apply this axiomatic system in any context where we have an available set of non-negative real numbers. This can be the case, for instance, when we dispose of non-negative coefficients (noted p) of a given set and target the estimation of the related model parameters through their reparametrization (Golan, Judge & Miller, 1996). Naturally, we will come back to such applications, and an estimation approach using probabilities and support space simultaneously will be presented. This underscores an important role to be assigned to the probability form of numbers, which motivated the selection of the axioms below. We will want our information measure $I(p)$ to have several properties:

1. Information is a non-negative quantity, i.e., $I(p) \geq 0$. Following what has been presented above on information definition (see 2.4), one may generalize this property to convexity in the next theorem:
   Theorem: $I(p_1 : p_2)$ is almost positive defined, that is $I(p_1 : p_2) \geq 0$ with equality if and only if $f_1(x) = f_2(x)$ $[\lambda]$.
   We will not demonstrate this theorem (see Kullback, 1959, pp. 14–15); we just provide the reader with the essence channelled through it. The above theorem explains that in the mean, discrimination information from statistical observations is positive. It follows from what has been previously said that no discrimination information will result if the distribution of observations is the same $[\lambda]$ under hypothesis one and two. A typical example—as we will see later—may constitute maximum entropy and cross-entropy principles. In that case, when non-informative consistency moments from observations are not provided, minimum cross-entropy declines into maximum entropy.
2. If an event has probability 1, certainty follows, and we get no information from the occurrence of the event: $I(p = 1) = 0$.
3. If two independent events occur (whose joint probability is the product of their individual probabilities), then the information we get from observing the events is the sum of the two pieces of information:
   $I(p_1 p_2) = I(p_1) + I(p_2)$. This property is referred to as additivity. Note that this property presents a valuable feature; it represents the basis of the logarithmic form of information. Intuitively, that means that a sample of n independent observations from the same population provides n times the mean information in a single observation.
   In the case of non-independent events, the additive property is retained, but in terms of conditional information.
4. Finally, as already stipulated in the preceding section, we will want our information measure to be a continuous (and, in fact, monotonic) function of the probability—slight changes in probability should result in slight changes in information. For consistency with the properties above, it can be useful to show the logarithmic feature of statistical information in the following way:

   1. $I(p^2) = I(pp) = I(p) + I(p) = 2I(p)$ (2.10)

   2. Through inductive reasoning, one can generalize (2.10) and write $I(pn) = nI(p)$

   3. $I(p) = I((p^{1/m})^m) = \text{m}(p^{1/m})$

   and we have

   $$I(p^{1/m}) = \frac{1}{m} I(P)$$

and, once again, we can generalize in the following way :

$$I(p^{n/m}) = \frac{n}{m} I(p)$$

4. The property of continuity allows us to write, for $0 < p \le 1$ and a real number α:

$$I(p^{\alpha}) = \alpha I(P).$$

From (2.10), one can observe that an operator transforming the probability p at the power $_{n/m}$, (that is, $p^{n/m}$) into an information measure $I(p^{n/m})$ displays a logarithmic property of additivity. This allows us to write a general, useful relation:

$$I(p\ ) = -\log_b(p) = \log_b(1/p) \text{ for base } b. \tag{2.11}$$

For other information properties not directly connected with the aim of this work, such as invariance or sufficiency, which will not be presented here, see Jaynes (1994), Kullback (1959). Furthermore, in the coming chapters, additional properties for different forms of entropy will be presented, such as concavity and stability (common for both Shannon-Gibbs and Tsallis entropies) or extensivity (common for both Shannon-Gibbs and Renyi (1961) entropies).

As a final remark of this section, it is important to note that the above logarithmic nature of information as explained in (2.11)—for the case of independent events—is limited to ergodic systems which convey additive-extensive properties of information in the case of independent events.

2 Ill-posed Inverse Problem Solution and the Maximum Entropy Principle

2.1 Introduction

As explained in the introduction, many economic relationships are characterized by indeterminacy. This may be because of long-range feedback and complex correlations between source and targets, thus rendering causal relationships more difficult to investigate.

In this part of the work, the formal definition of the inverse problem will be discussed. A Moore-Penrose approach will be presented for solving this kind of problem and its limits will be stressed. The next step will be to present the concept of the maximum entropy principle in the context of the Gibbs-Shannon model. Extensions of the model by Jaynes and Kullback-Leibler will be presented and a generalisation of the model will be implemented to take into account random disturbance. The next step will concern the non-ergodic form of entropy known in the literature of thermodynamics as non-extensive entropy or non-additive statistics. There will be a focus on Tsallis entropy, and its main properties will be presented in the context of information theory. To establish a footing in the context of real world problems, non-extensive entropy will be generalized and then random disturbances will be introduced into the model. This part of the work will be concluded with the proposition of a statistical inference in the context of information theory.

2.2 The Inverse Problem and Socio-Economic Phenomena

An inverse problem (e.g., Thikonov et al. (1977), Bwanakare (2015), Golan et al. (1996)) explains a situation where one tries to capture the causes of phenomena for which experimental observations represent the effect.

The essence of the inverse problem is conveyed by the expression:

$$Y = X\beta + \xi \tag{2.12}$$

or its equivalent in continuous form:

$$Y(\zeta) = \int_D g(X)B(X,\zeta)dX + b(\zeta) \tag{2.13}$$

where
X represents the state space,
Y designates the observation space,
D defines the Hilbert support space of the model,
B is the transformation kernel linking measures X and Y,
$b(\zeta)$ displays random error process.

In classical econometrics, when given a state X, an operator B and, as happens most of the time, a disturbance term (ζ), what is Y? This is referred to as a forward problem. In social science, one must often cope with the above random (Gaussian or not) disturbance term, and this usually complicates matters in spite of significant, recent developments in econometrics, particularly concerning stochastic time-series analysis (Engle & Granger, 1987). Furthermore, the inverse question is more profound: Given y and a specific B, what is the true state X?

If B should also be a functional of X, the problem becomes arbitrarily complex. Correlation between (ζ) and X will be at the base of such additional complexity.

Every day, psychologists cope with such inferential problems. Patients display identical symptoms from different sicknesses. Health practitioners need more historical (*a priori*) information on patients to try to find the solution.

In economics, the same national output growth rate may result from different combinations of factors. One of the main problems encountered by practicing economists is isolating the causes of economic phenomena once they have occurred. In most cases, the economist becomes inventive in finding an appropriate hypothesis before trying to solve the problem. As an example, in the case of a recession or financial turbulence, it is usually difficult to point to principal causes and fix them. Schools of economics suggest different, even contradictory, solutions—the legacy of its inverse problem nature.

In empirical research, many techniques exist to try to solve the inverse problem. In the context of the present work, the presentation will be limited to those more applicable to matrix inversion, like the Moore-Penrose pseudo-inverse approach, and, naturally, maximum entropy based approaches. The approach better known in economics for updating national accounts on the basis of bi-proportionalities will then be added to these two techniques.

2.2.1 Moore-Penrose Pseudo-Inverse

Let us consider the discrete and determinist case and rewrite (2.12) as follows:

$$Y = XB = X\rho \tag{2.14}$$

In the right equality ρ reflects the case where we have to deal with a ratio or probability parameter, for example, after reparametrizing B. We then have:

$$\begin{aligned} &\rho = YS \Leftrightarrow Y = XBS \\ &\hat{\rho} = BY \Leftrightarrow Y = XYV \\ &Y = X\hat{\rho} = XVY = XBX\rho, \end{aligned} \tag{2.15}$$

which means:
$XBX = X$ and V, representing the generalized inverse matrix (Golan, 1996), (Kalman, 1960).

This is a matrix with the symbol $B+$ that satisfies the following requirements:

$B B^+ B = B$,
$B^+B B^+ = B$,
B^+B is symmetric,
$B B^+$ is symmetric.

Following Theil (1967), a unique B^+ can be found for any matrix: square, non-singular or not. When the matrix B^+ happens to simultaneously be square and non-singular, then the generalized inverse will be the ordinary inverse B^-. The problem that interests us is the over-determined system of equations

$Y = XB$

where B has n rows, $K < n$ columns and column rank equal to $R \leq K$.

If we retain the particular case when R equals K to ensure the existence of $(B'B)^{-1}$, then the generalised inverse of B is

$B^+ = (B'B)^{-1}B'$

as can be easily verified. A solution to the system of equations can be presented as:

$X = B^+ Y$.

Following Green (2003, p. 833), we note in this case that the length of this vector minimizes the distance between Y and BX, according to the least squares properties method. This distance will naturally remain equal to zero if y lies in the column space of B.

If we now retain the more general case where B does not have full rank, the above solution is no longer valid and a spectral decomposition using the reciprocals of the characteristic roots is involved to compute the inverse which becomes:

$B^+ = C_1 A_1^{-1} C_1' B'$

where C_1 are the R characteristic vectors corresponding to the non-zero roots arrayed in the diagonal matrix A_1.

The next and last case is the one where B is symmetric and singular, that is, with the rank $R \leq K$. In such a case, Moore-Penrose inverse is computed as in the preceding case but without pre-multiplying by B'. Thus, for such a symmetric matrix,

$$B^+ = C_1 A_1^{-1} C_1', \tag{2.16}$$

with A_1^{-1} being a diagonal matrix of the reciprocals of the non-zero roots of B.

It is important to note that only matrix B with full rank ensures a minimum distance between Y and BX. In other cases, there may exist an infinite number of combinations of elements of matrix B or $\hat{\rho}$ which satisfy (2.14).

To conclude, in spite of strong advantages of the Moore-Penrose generalised inverse, outputs will not always reflect an optimal solution.

2.2.2 The Gibbs-Shannon Maximum Entropy Principle and the Inverse Problem

Let us introduce the concept of Shannon entropy by continuing with the case of pure linear inverse problem solution discussed above. The simplest (one dimensional case) example is the Jaynes dice inverse problem.

If a die is fair, and we throw it a large number of times n, with k different output modalities[9] ($k = 1,\dots, K$), the expected value will be 3.5, as from a uniform distribution with probability f_k equal 1/6. How can one infer about p_k if we have 'loaded' (unfair) dice and the expected value of the trial becomes:

$$4.5 = \sum_{k=1}^{K} k p_k \qquad (2.17)$$

where frequencies p_k is n_k/n ?

In this case, the central question is: Which estimate of the set of frequencies would most likely yield this number? The problem is underdetermined since there are many sets of f_k that can be found to fit the single datum of Equation (2.17). Here we have to deal with a multinomial distribution where the multinomial coefficient W is given by:

$$W = \frac{\mathrm{N}!}{\mathrm{N}p_1!\mathrm{N}p_2!..\mathrm{N}p_k!} = \frac{\mathrm{N}!}{\mathrm{N}k_1!\mathrm{N}k_2!..\mathrm{N}k_k!}$$

Deriving and using the Stirling approximation $\ln x! \cong x\ln x - x$ for a large number of N, we get the Shannon entropy formulation:

$$Max_p H(p) = -\sum_{k=1}^{K} p_k \ln p_k \qquad (2.18)$$[10]

In the case of a die, parameter K equals 6, and W is the multinomial coefficient, i.e., the number yielding a particular set of frequencies among 6^N possible outcomes.

9 Generally, if the number of trials is equal to n, we will have n_k possible outputs corresponding to each modality k with $n = \sum_k n_k$. Thus, the frequency $p_k = n_k/n$ is related to each modality k.

10 Note that the generalized form of Shannon entropy in the continuous case has the form: $Max_{f(y)} H(f(y)) = -\int f(y) \log f(y) dy$.

We need only find the set of frequencies maximizing W in order to find the set that can be realized in the greatest number of ways. This is the most plausible combination in the case of fair dice.

This turns out to convey the same logic as maximizing Shannon Gibbs entropy. Thus, starting from two pieces of information, that is, the number k equal to six and N, a large number of trials, we are able to derive six probabilities related to a die distribution.

Next, Jaynes (1994) maximized the Shannon function through the restriction of consistent information at hand. This opened entropy theory application to many scientific fields, including the social sciences.

Thus, if we add to the formulation (2.18) the moment-consistency and the adding up-normalization constraints, we then get:

$$Max_p H(p) = -\sum_{k=1}^{K} p_k \ln p_k \tag{2.19}$$

subject to:

$$\sum_{k=1}^{K} p_k f_t(x_k) = y_t, 1 \le t \le T \tag{2.20}$$

$$\sum_{k=1}^{K} p_k = 1 \tag{2.21}$$

where $\{y_1, y_2, \ldots, y_t\}$ denotes a set of observations (e.g., aggregate accounts or their averages) being consistent with a function $f_t(x_k)$ of explicative variables weighted by a corresponding distribution of probabilities $\{p_1, p_2, \ldots, p_k\}$. As usually happens, T is less than K, and the problem is ill-posed (underdetermined).

Two main results emerge from the above formulation. First, if all events are independent or quasi-independent (locally dependent) and equally probable, then the above entropy is *a linear function of the number of the possible system states* and then is *extensive*[11].

A second fundamental result is connected with information theory and suggests that a Gaussian variable has the largest entropy among all random variables of equal variance (see Papoulis, 1991 for proof). In the next chapter on non-extensive entropy, a measure to assess the divergence of a given distribution from Gaussian distribution will be presented.

11 For this reason, as earlier alluded to, the Gibbs-Shannon entropy is called *extensive*. In reverse, as it will be commented on in the coming sections, the hypothesis of long-range correlation between events leads to the concept of *non-extensive* entropy (e.g., Tsallis entropy) suggesting an entropy no longer being a linear function of data.

Table 1: Recovering probability distribution of an unbalanced die through the maximum entropy principle.

P1	***P2***	***P3***	***P4***	***P5***	***P6***
0.054	0.079	0.114	0.166	0.240	0.348

Coming back to the dice case, maximization of Shannon entropy in (2.19), that is $H(P) = -P' \ln P$ under Jaynes consistency, leads to the distribution presented in Table 1. To solve this inverse problem of six unknowns, the only two pieces of information available are the expected value—from the experiments in this example—assumed to be equal to 4.5 and the information that the probability of different possibilities adds up to one. However, since we are dealing with unbalanced dice, we have no idea about the distribution.

The next chapters extend the Shannon-Gibbs-Jaynes maximum entropy principle with Kullback-Leibler relative entropy. The next to the last targeted presentation will deal with the general linear entropy model, that is, the one with a stochastic component. To conclude, Tsallis power law distribution to generalize Kullback-Leibler cross-entropy will be considered.

2.2.3 Kullback-Leibler Cross-Entropy

Kullback (1959), Good (1963) extended the Jaynes-Shannon-Gibbs model by formulating the principle of minimum (cross or relative) entropy. Using an *a priori* piece of information q about unknown parameter p, the resulting formulation is as follows:

$$Min\_H(p,q) = \sum_{k=1}^{K} p_k \ln(p_k / q_k) = p'\ln p - p'\ln q \tag{2.22}$$

under restrictions:

$$Y = XP \tag{2.23}$$

$$P'1 = 1 \tag{2.24}$$

where p = $(p1, \ldots, p_K)'$, $q = (q1, \ldots, q_K)$.

These restrictions are the same as those presented earlier. In the criterion function (2.22), *a posteriori* and *a priori* vectors or matrices p and q are confronted with the purpose of measuring entropy reduction resulting from exclusive new content of data information.

One should note that when q is fully consistent with moments, then $p = q$ and the distribution becomes uniform with $q_k = 1/K$. This leads to the solution of the maximum entropy principle.

Thus, the cross-entropy principle stands for a certain form of generalization of maximum entropy. Relation (2.22) above is an illustration of the previous Kullback formulation in (2.8) as a mean information from (2.23) and (2.24) for discrimination in favour of p against q.

2.3 General Linear Entropy Econometrics

In social science, it is rare to encounter the situation described by the relation (2.14) where the random term is meaningless as is often encountered in the experimental sciences. Social phenomena are particularly affected by stochastic components. Let us rewrite it below in its generalized form:

$$y_i = \sum_{j-1}^{K} B_j X_j + \zeta_i \qquad (2.12')$$

with the random term $\zeta_i \in e$ and
$i = (1,\ldots, I)$ (I being the number of observations); K is the number of model parameters to be estimated.

2.3.1 Reparametrization of Parameters

Following Golan et al. (1996), we first reparametrize the above generalized entropy model (2.12').

We treat each B_j (j = 1,..., K) as a discrete random variable within a compact support and 2 < M < ∞ possible outcomes. So, we can express B_j as:

$$B_k = \sum_{m=1}^{M} p_{km} v_{km} \quad \forall k \in K \qquad (2.25)$$

where p_{km} is the probability of outcome v_{km} and the probabilities must be non-negative and sum up to one.

Similarly, let us treat each element ζ_i of e as a finite and discrete random variable with compact support and 2 < M < ∞ possible outcomes centred on zero. We can express ζ_i as:

$$\zeta_i = \sum_{j=1}^{J} r_{nj}.w_{nj} \qquad (2.26)$$

where r_n is the probability of outcome w_n. The term ζ_i, like any prior value in the model, reflects Bayesian properties and is not a fixed value as in the case of classical econometric models. In practice, support sets with three or more points[12] are used to take into account higher moments of the distribution during the process of information recovery.

2.4 Tsallis Entropy and Main Properties

2.4.1 Definition and Shannon-Tsallis Entropy Relationships

This relatively new form of entropy is emerging over an immense area of applications in social science, including economics. One of the fields of interest is modelling and predicting markets of financial returns (Drożdż & Kwapień, 2012), (Grech & Pamula, 2013). Nevertheless, due to the high frequency nature of Big Data in Official Statistics (e.g., Braaksma & Zeelenberg, 2015), the PL-based non-extensive entropy econometrics should be seen as a potential and natural estimation device in this new statistical area. As in statistical physics, socioeconomic random events display two types of stochastic behaviour: ergodic and non-ergodic systems. Whenever isolated in a closed space, ergodic systems dynamically visit with equal probability all the allowed micro-states (Gell-Mann & Tsallis, 2004). However, it seems logical to imagine systems visiting the allowed micro-states in a much more complex way than defined by ergodicity. The financial market is a well-known example of such complex systems, as characterized by multifractal dimensions (Drożdż & Kwapień, 2012), (Grech & Pamula, 2013). Other examples include income distribution inside a given region, evolution of a given disease inside a region, size of cities, or cellular structure. These forms seem to display an organized structure owing to long-range correlation between micro-elements, heavy queues with respect to Gaussian distribution, scale-invariant structures, and criticality. Such phenomena would be better described by a stable law-based Levy process, like power law distribution.

Shannon-Kullback-Leibleir Equations (2.22–2.24) are generalized by Tsallis relative entropy formulation. To emphasize consistency among the principal formulations, it is worthwhile to reiterate the statistical theory connection between the above relations and the Kullback relation presented in (2.8) or to some extent (2.9), which

12 Golan, Judge, and Miller (1996) suggest the Chebyshev inequality as a good starting point to define the error support set: $\Pr[|x| < v\sigma] \geq v^{-2}$ where v is a positive real and x a random variable, such that $E(x) = 0$ while $\mathrm{var}(x) = \sigma^2$. This inequality leads to the three-sigma rule (Pukelsheim, 1994) for $v = 3$, i.e., to the probability $\Pr[-3\sigma < x < 3\sigma]$, which is at least 0.88 and higher when x displays a standard normal distribution. Let us remember that this inequality has the additional advantage of being independent of distribution laws.

measures the divergence between two hypotheses H_1 and H_2. A similar concept will be introduced in the case of non-extensive entropy, which will constitute the final step of Shannon entropy extensions.

Let us generalize the Shannon Gibbs inverse problem through ordinary differential equation characterization (Tsallis, 2009). First, we need to introduce the three simplest—in terms of dynamic complexity—differential equations and their inverse functions,

$$\frac{dy}{dx} = 0 \quad (y(0) = 1). \tag{2.27}$$

Its solution is y=1 ($\forall x$), and its inverse function is X=1 ($\forall y$).

The next simplest differential equation is

$$\frac{dy}{dx} = 1 \quad (y(0) = 1). \tag{2.28}$$

Its solution is y=(1 + x) and its inverse Y=(x – 1).

The next higher step in increasing complexity is the differential equation

$$\frac{dy}{dx} = y \quad (y(0)=1). \tag{2.29}$$

Its solution is $y = e^x$, and its inverse is $y = \ln x$.

Note that the latter inverse equation satisfies the additive property:

$$\ln(x_a x_b) = \ln(x_a) + \ln(x_b). \tag{2.30}$$

Following Gell-Mann & Tsallis (2004) and trying to unify the three cases (without preserving linearity), we get:

$$\frac{dy}{dx} = y^q \;\; (y(0)=1;\, q \in \Re). \tag{2.31a}$$

We observe that this expression displays power-law distribution form.
Its solution is

$$y = [1 + (1 - q)x]^{\frac{1}{1-q}} \equiv e_q^x (e_1^x = e^x),$$

and its inverse function is

$$y = \frac{x^{(1-q)} - 1}{1-q} \equiv \ln_q x \; (\ln_1 x = \ln x). \tag{2.31b}$$

The above represents the non-extensive (Tsallis) entropy formula. Though it will be discussed in the next section, let us immediately show here the relationship between Shannon and Tsallis entropies through the next pseudo-additive property:

$$\ln_q(x_a x_b) = \ln(x_a) + \ln(x_b) + (1 - q)\ln_q(x_a)\ln_q(x_b) \tag{2.32}$$

for $q \to -\infty$, $q = 0$, $q = 1$ we obtain the three initial cases (2.27 – 2.29), respectively.

In particular for $q = 1$, we then obtain (after using l'Hôpital's rule) the solution of (2.29), the case of Shannon Gibbs entropy. The expression (2.30) states that if two systems x_a and x_b are logarithmically multiplied, the output is the additive sum of these systems in a logarithmic scale. This explains why Shannon entropy is sometimes referred to as additive entropy. This observation has been taken from (2.21) to emphasize that Shannon entropy is a direct function of data. The term q is referred to as "q-Tsallis." When it is equal to unity, we reach in this limiting case the Shannon entropy.

Tsallis entropy should now be described and compared with other entropy forms. This description indirectly replies to the question of why Tsallis or Shannon entropy rather than Renyi entropy or another is appropriate for a given problem.

2.4.2 Characterization of Non-Extensive Entropy

2.4.2.1 Correlations

Following Tsallis (2009), suppose we have a system composed of subsystems[13] A (with W_A possibilities of complexities) and B (with W_B possibilities of complexity). Their joint probabilities can be presented as $\left\{p_{ij}^{AB}\right\}$ ($i = 1,2,\ldots, W_A$ $j = 1,2,\ldots, W_B$) and marginal probabilities as

$$p_i^A \equiv \sum_{j=1}^{W_B} p_{ij}^{AB} \text{(hence } \sum_{i=1}^{W_A} p_i^A = 1) \text{ and}$$

$$p_j^B \equiv \sum_{i=1}^{W_A} p_{ij}^{AB} \text{(hence } \sum_{j=1}^{W_B} p_j^B = 1).$$

In general,

$$p_i^A p_j^B \neq p_{ij}^{AB} \tag{2.33}$$

if they happen to be equal, then A and B are said to be probabilistically independent. Otherwise, they are *dependent* or *correlated*. Let us then define entropies:

$$S_q\left(p^A\right) \equiv S_q\left(\left\{p_i^A\right\}\right)$$

and

$$S_q\left(p^B\right) \equiv S_q\left(\left\{p_i^B\right\}\right).$$

13 For models to be presented later, subsystem A can be considered as a data generating process and B as a subsystem of disturbances. After reparametrization, these two subsystems will be associated in terms of probabilities.

More interestingly, the conditional entropies definition, that is,

$S_q(p_{A/B})$ and $S_q(p_{B/A})$,

may deserve closer attention, as it can intervene for the definition of estimation precision of a model whether or not the hypothesis of independence between the model variables and its random terms has been accepted. If and only if A and B are independent,

$$S_q(p_{A/B}) = S_q(p^A)$$

and

$$S_q(p_{B/A}) = S_q(p^B).$$

Next,

$$S_q(AB) \equiv S_q(\{p_{ij}^{AB}\}),$$

in general, satisfies:

$$S_q(AB) = S_q(p^A) + S_q(p_{B/A}) + (1-q)S_q(p^A)S_q(p_{B/A}) = S_q(p^B) + S_q(p_{A/B}) + \\ + (1-q)S_q(p^B)S_q(p_{A/B}) \quad (2.34)$$

Finally, to be more explicit than in the previous section, S_q is said to be non-extensive in the sense that given two independent random systems A and B, i.e., $P(A, B) = P(A)P(B)$, then,

$$S_q(AB) = S_q(p^A) + S_q(p^B) + (1-q)S_q(p^A)S_q(p^B) \quad (2.35)$$

In the next, empirical part of this book, for inferential purposes and for optimal simplification of numerical computations, this formula will play a key role in determining the level of entropy of a complex system under the hypothesis of independence of subsystems, i.e., between the model and the random term.

2.4.2.2 Concavity

The concept of concavity is important since, among others things, it allows us to determine whether or not a system is stable. Stability is a meaningful concept in econometrics since it implies stationarity of a process in a given system. Testing for stationarity and cointegration using entropy distribution seems thus to be an open area of further research[14].

14 However, the job may be rendered difficult since optimal equilibrium responding to economic

S_q is concave (convex) for all probability distributions and all $q \succ 0 (q \prec 0)$ (Gell-Mann & Tsallis, 2004). Let us follow the traditional mathematical definition of concavity and let {*pi*} and $\{p_i'\}(i = 1,2,\ldots,W)$ be two arbitrary probability distributions. The next relation of intermediate distribution follows:

$$p_i'' \equiv \lambda p_i + (1-\lambda) p_i' (0 \prec \lambda \prec 1).$$

By concavity we mean that it can be proven that for all λ,

$$S_q\left(\left\{p_i''\right\}\right) \geq \lambda S_q(\{p_i\}) + (1-\lambda) S_q\left(\left\{p_i'\right\}\right).$$

2.4.3 Tsallis Entropy and Other Forms of Entropy

Let us first review the mathematical main forms of entropies before presenting their most important distinctive properties.

$$S_{BG}(p) \equiv -\sum_i pi \ln pi = S_1 = S_1^R = S_1^{LVRA} \tag{2.36}$$

$$S_q(p) \equiv \left(1 - \sum_i p_i^q / (q-1)\right) \tag{2.37}$$

$$S_q^R(p) \equiv \left(\ln \sum_i p_i^q\right) / (1-q) = \ln\left[1 + (1-q) S_q\right] / (1-q) \tag{2.38}$$

$$S_q^{LVRA}(p) \equiv (S_q) / \left(\sum_i p_i^q\right) = \left(1 - \left[\sum_i p_i^q\right]^{-1}\right) / (1-q) = S_q / \left[1 + (1-q) S_q\right] \tag{2.39}$$

A key element deserves attention here. We see from the first mathematical relation in (2.36) above that Shannon-Gibbs entropy may be generalized, too, by Renyi entropy (2.38) or by the normalized non-extensive form (2.39), independently introduced by Landsberg & Vedral (1998) and by Rajagopal and Abe (2000). Both forms of entropy are monotonically increasing functions of Sq. Tsallis (Gell-Mann & Tsallis, 2004, p. 11) poses and explains a relevant question concerning relationships between these forms of entropy. In fact, after pointing out that monotonicity *makes Sq, SqR, and SqN* extreme for the same probability distribution, he asks why not base thermodynamics on *SqR* or *SqN* rather than only on Tsallis entropy. The response lies in the

---

laws does not necessarily fit into optimal entropy equilibrium. This problem will be briefly covered, later.

disadvantages of these two forms of competitive entropy. In fact, it happens that they are not concave for all positive values of q, but only for $0 \prec q \leq 1$. Since many physically meaningful phenomena for which q are higher than unity exist, this becomes a serious drawback of both competitive entropies. As far as economic, financial, or social phenomena are concerned, the problem does not allow for any ambiguity since, as we will see in the next section, $1 \leq q \prec 5/3$. For the majority of them[15], extreme events are on average more frequent (with persistence) than predicted by Gaussian law and not the reverse (i.e., less frequent—with persistence—than predicted by Gaussian law). Tsallis entropy thus remains the one form that not only generalizes SG entropy but also ensures concavity (stability) inside the whole finite interval where probability distribution is defined. The reader should thus far understand why non-extensive Tsallis entropy has been recently used to generalize all other forms of entropy, at least in many fields where entropy is applied.

2.4.3.1 Characterization

In the following table, we illustrate different links between the commonly used forms of entropy with respect to the characterization in Table 2. "Yes" and "No" correspond, respectively, to what, according to recent thermodynamics literature (Gell-Mann & Tsallis, 2004), are thermodynamically allowed and forbidden violations of the Boltzmann–Gibbs (BG) entropy properties.

Table 2: Comparison of different forms of entropy with regard to important properties

Property	Entropy			
	S_{BC}	S_q	S_q^R	S_q^{LVRA}
Extensive $(\forall q)(p_{ij}^{AB} = p_i^A p_j^B)$	Yes	No	Yes	No
Concave $(\forall q \succ 0)$	Yes	Yes	No	No
Stable $(\forall q \succ 0)$	Yes	Yes	No	No
Optimizing distribution $(\forall q)$	exponential Law	Power Law	Power Law	Power Law

Source: own, based on Tsallis (2009) and Gell-Mann & Tsallis (2004)

15 For example, for stock market returns, q is around 1.4, far enough from the unity which characterises Gaussian distribution.

NB: R stands for Renyi, and $S_q^{LVRA} \equiv S_q^N$ (LVRA and N stand for Landsberg-Vedral-Rajagopal-Abe and normalized, respectively (Gell-Mann & Tsallis, 2004).

2.4.3.2 Scale of q-Tsallis Index and its Interpretation

Following the thermodynamcs literature built on Lévy-like anomalous diffusion, it has been shown that $p(x) \propto e_q^{-\beta x^2}$ optimizes

$$S_q = \frac{1 - \int dx[p(x)]^q}{q-1}$$

under appropriate constraints. If one convolutes n times $p(x)(n \to \infty)$, we approach a Gaussian distribution if $q \prec 5/3$, and a Lévy $L_{\gamma L}(x)$ if $5/3 \prec q \prec 3$. The index γL of Lévy distribution is related to q as follows:

$$q = \frac{\gamma L + 3.}{\gamma L + 1} \quad (5/3 \prec q \prec 3).$$

Thus, in empirical applications, the value of q should vary inside an interval from unity to 5/3, which corresponds to cases of finite variance for phenomena dwelling within the Gaussian basin of attraction.

2.5 Kullback-Leibler-Tsallis Cross-Entropy

2.5.1 The q-Generalization of the Kullback-Leibler Relative Entropy

Kullback-Leibler-Tsallis cross-entropy is known in literature as the q-generalization of Kullback-Leibler relative entropy. The Kullback-Leibler-Shannon entropy introduced in Part II can be q-generalized (Tsallis, 2009) in a straightforward manner. The discrete version becomes:

$$I_q(p, p_o) \equiv \sum p_i \frac{[p_i / p_{oi}]^{q-1} - 1}{q-1} \tag{2.40}$$ [16]

since with any real $r \succ 0$, one has the following properties:

$$\frac{r^{q-1} - 1}{q-1} \geq 1 - \frac{1}{r} \quad \text{if } q \succ 0 \tag{2.41}$$

16 In a continuous case, we have:

$$I_q(p, p^{(0)}) \equiv -\int dx p(x) \ln_q \left[\frac{p^{(0)}(x)}{p(x)} \right] = \int dx p(x) \frac{[p(x)/p^{(0)}(x)]^{q-1} - 1}{q-1}$$

$$= 1 - \frac{1}{r} \quad \text{if } q = 0$$

$$\leq 1 - \frac{1}{r} \quad \text{if } q \prec 0$$

Thus, retaining the practical case of $q \succ 0$, we can write:

$$\frac{\left[p(x)/p_o(x)\right]^{q-1} - 1}{q-1} \geq 1 - \frac{p_o(x)}{p(x)}.$$

Hence[17],

$$\sum p(x) \frac{\left[\dfrac{p(x)}{p_o(x)}\right]^{q-1} - 1}{q-1} \geq \sum p(x)\left[1 - \frac{p_o(x)}{p(x)}\right] = 1 - 1 = 0.$$

Therefore, coming back again to the generalized K-Ld cross-entropy, we have[18]:

$$\begin{aligned} I_q(p,p_o) &\geq 0 \text{ if } q \succ 0, \\ &= 0 \text{ if } q = 0, \\ &\leq 0 \text{ if } q \prec 0. \end{aligned} \tag{2.42}$$

Thus, as Tsallis (2009) has made us aware, the above q-Kullback-Leibler index has the same basic property as the standard Kullback-Leibler entropy and can be used for the same purpose while having the additional advantage of an adaptive q according to the system with which we are dealing.

There exist two different versions of the Kullback-Leibler divergence (K-Ld) in Tsallis statistics, the usual generalized K–Ld shown above and the generalized Bregman K–Ld. According to Venkatesanet et al. (Plastino & Venkatesan, 2011), problems have been encountered in empirical thermodynamics trying to reconcile these two versions. Unfortunately—or fortunately!—the same problems seem to reappear while applying this theory in social science since every version of generalized K-Ld leads to different outputs. Let us try to synthesize what recent literature says about this problem.

---

17 It is straightforward to derive this property in the case of the continuous case.

18 The same conclusion is obtained by using Jensen's inequality (e.g., Gell-Mann & Tsallis, 2004).

2.5.2 Tsallis Versions of the Kullback-Leibler Divergence in Constraining Problems

This short section represents the final bridge between theory and the applications in the last parts of this work. In a recent study, Plastino & Venkatesan (2011) lay out interesting aspects of empirical research when q-generalized K-Ld cross-entropy is associated with constraining information. Since, in the social sciences, we particularly need discrete forms of these relative entropies, let us first rewrite these forms before commenting on their conditions of applicability:

$$I_q(p,p_o) \equiv \sum p_i \frac{[p_i/p_{oi}]^{q-1}-1}{q-1} \tag{2.43}$$

$$I_q[p\|p_0] = \frac{1}{q-1}\sum_i p_i\left[(p_i)^{q-1}-(p_{0i})^{q-1}\right] - \sum_i (p_i - p_{0i})(p_{0i})^{q-1} \tag{2.44}$$

The form (2.43) is the one derived directly from Kullback-Leibler formalism and presented in (2.40). The second form is referred to as the generalized Bregman form of K-Ld cross-entropy, and it is more appealing than (2.43) from an information-geometric viewpoint (Plastino & Venkatesan, 2011) even if it does contain certain inherent drawbacks.

A study by Abe and Bagci (2005) has demonstrated that the generalized K–Ld defined by (2.44) is jointly convex in terms of both p_i and p_{0i} while the form defined by (2.43) is convex only in terms of p_i. A further distinction between the two forms of the generalized K–Ld concerns the property of composability. While the form defined by (2.44) is composable, the form defined by (2.43) does not exhibit this property.

The second interesting aspect for practitioners concerns the manner in which mean values are computed. Non-extensive statistics has employed a number of forms in which expectations may be defined. The first among these are the linear constraints initially used by Tsallis (2009), also known as normal averages, that is:

$$\langle y \rangle = \sum_i p_i y_i$$

The second is the Curado-Tsallis (C-T) constraints of the form:

$$\langle y_q \rangle = \sum_i p_i^q y_i$$

and the normalized Tsallis-Mendes-Plastino (TMP) constraints (also known as q-averages or an escort distribution) of the form:

$$\langle y_q \rangle = \sum_i \frac{p_i^q}{\sum_i p_i^q} y_i$$

A fourth—less applied by practitioners—constraining procedure is the optimal Lagrange multiplier approach.

Among these four methods to describe expectations, the most commonly employed by Tsallis practitioners is TMP, referred to as escort distribution.

Recent work by Abe (2009) suggest that, in generalized statistics, expectations defined in terms of normal averages, in contrast to those defined by q-averages, seem to display higher consistency in material chaos hypotheses. Recent reformulation of the variational perturbation approximations in non-extensive statistical physics followed from these findings. To my knowledge, application in the social sciences to assess the universality of this finding has not been done yet.

Finally, there is the issue of consistency. This stems from the form of the generalized K–Ld defined by (2.43) being consistent with expectations and constraints defined by q-averages ("prominently" the TMP) while, on the other hand, the generalized Bregman K–Ld defined by (2.44) is consistent with expectations defined by normal averages.

Thus, through reformulations of an empirical inverse problem, this last point may play a key role since non-appropriated constraints should lead to a non-optimal solution in the best case or to computational problems, as is often the case.

2.6 A Generalized Linear Non-Extensive Entropy Econometric Model

2.6.1 A General Model

This section presents a generalized linear non-extensive entropy econometric approach to estimate econometric models. Following Golan et al. (1996), we first reparametrize the generalized linear model of the equation (2.12') rewritten below:

$$y_i = \sum_{k=1}^{K} B_k X_k + \zeta_i \tag{2.12'}$$

with, once again, the random term $\zeta_i \in e$ and $i = (1,..., I)$ (I being the number of observations); K is the number of model parameters to be estimated; where B values are not necessarily constrained between 0 and 1, and ζ is an unobservable disturbance term with finite variance, owing to the nature of economic data that exhibits error observation from empirical measurement or random shocks. If we treat each B_j (k = 1... K) as a discrete random variable with compact support and $2 < M < \infty$ possible outcomes, we can express B as:

$$B_k = \sum_{m=1}^{M} p_{km} v_{km}, \forall k \in \{1, \ldots, K\} \tag{2.45}$$

where p_{km} is the probability of the outcome v_{km}. The probabilities must be non-negative and add up to one. Similarly, by treating each element ζ_i of ζ as a finite and discrete random variable with compact support and $2 < M < \infty$ possible outcomes centred around zero, we can express ζ_i as:

$$\zeta_i = \sum_{j=1}^{J} r_{ij} w_{ij} \tag{2.46}$$

where r_i is the probability of outcome w_i on the support space j, with $j \in \{1,\ldots,J\}$ and $i \in \{i = 1,\ldots,N\}$. Note that the term e (an estimator of ζ) can be fixed as a percentage of the explained variable, as an *a priori* Bayesian hypothesis. Posterior probabilities within the support space may display non-Gaussian distribution. The element v_{km} constitutes *a priori* information provided by the researcher while p_{km} is an unknown probability whose value must be determined by solving a maximum entropy problem. In matrix notation, let us rewrite $\beta = V \cdot P$ with $p_{km} \geq 0$ and

$$\sum_{k=1}^{K} \sum_{m>2\ldots M} p_{km} = 1,$$

where again, K is the number of parameters to be estimated and M the number of data points in the support space. Also, let $e = r \cdot w$, with $r_{ij} \geq 0$ and

$$\sum_{i=1}^{N} \sum_{j>2\ldots J}^{J} r_{ij} = 1$$

for N the number of observations and J the number of data points on the support space for the error term. Then, the maximum Tsallis Entropy Econometric (MTEE) estimator can be stated as:

$$\max\left[H_q(p;r)\right] = \left\{\left[1 - \sum_k \sum_m \alpha \cdot (p_{km})^q\right] + \left[1 - \sum_i \sum_j (1-\alpha) \cdot (r_{ij})^q\right]\right\} \cdot (q-1)^{-1} \tag{2.47}$$

subject to

$$y_i = \sum_{k=1}^{K} B_k X_k + e_i = X \cdot \sum_{m=1}^{M} v_m \left(\frac{p_m{}^q}{\sum_{m=1}^{M} p_m{}^q} \right) + \sum_{j=1}^{J} w_j \left(\frac{r_j{}^q}{\sum_{j=1}^{J} r_j{}^q} \right) \tag{2.48}$$

$$\sum_{k=1}^{K} \sum_{m>2}^{M} p_{km} = 1 \tag{2.49}$$

$$\sum_{i=1}^{N} \sum_{j>2}^{J} r_{ij} = 1 \tag{2.50}$$

where the real q, as previously stated, stands for the Tsallis parameter.

Above, $H_q(p,r)$ weighted by α dual criterion function is nonlinear and measures the entropy in the model. The estimates of the parameters and residual are sensitive to the length and position of support intervals of β parameters. When parameters of the proposed mode[19] concern elasticity or error correct coefficients, the values of which lie between 0 and 1, then the support space should be defined inside the interval zero and one. In other cases, the support space may be defined between minus and plus infinity, according to the intuitive evaluation of the modeller. Additionally, within the same interval support, the model estimates and their variances should be affected by the number of support values (Golan et al., 1996). Increasing the number of point values inside the support space leads to improving the *a priori* information about the system. A few years of modelling with the maximum entropy approach seem to show that a well-defined support space is crucial to obtaining better results. The weights α and $(1 - \alpha)$ are introduced into the above dual objective function. The first term "of precision" accounts for deviations of the estimated parameters from the prior (defined under support space). The second, "prediction ex post," accounts for an empirical error term as a difference between predicted and observed data values of the model.

2.6.2 Parameter Confidence Interval Area

In this section, we will propose the normalized Tsallis entropy coefficient $S(\hat{a}_k)$ as an equivalent to a standard error measure in the case of classical econometrics. An equivalent of the determination coefficient R^2 will be introduced, also under the entropy symbol $S(\widehat{\Pr})$. The departure point is that the maximum level of entropy-uncertainty is reached when significant information-moment constraints are not enforced. This leads to a uniform distribution of probabilities over the k states of the system. As we add each piece of informative data in the form of a constraint, a departure from the uniform distribution will result, which means a lowering in uncertainty. Thus, the value of the proposed $S(\widehat{\Pr})$ below reflects a global departure from the maximum uncertainty for the whole model. Without giving superfluous theoretical details, we follow formulations in, e.g., Bwanakare (2014) and propose a normalized non-extensive entropy measure of $S(\hat{a}_k)$ and $S(\widehat{\Pr})$.

From the Tsallis entropy definition, S_q vanishes (for all q) in the case of $M = 1$; for M > 1, q > 0, whenever one of the $p_i (i = 1..M)$ occurrences equals unity, the remaining probabilities, of course, vanish. We get a global, absolute maximum of S_q (for all q) in the case of a uniform distribution, i.e., when all $p_i = {}^1/_M$. Note that we are interested,

19 As already presented, the expression $P_m = \frac{p_m^{\ q}}{\sum_{m=1}^{M} p_m^{\ q}}$ is referred to as *escort probabilities*, and we have for q=1 (then P_m is normalized to unity), that is, in the case of Gaussian distribution (Gell-Mann & Tsallis, 2004), (Tsallis, 2009).

for our economic analysis, in q values lying inside the interval (1, 5/3). In such an instance, we have for our two systems:

$$S_q(p) = (M^{1-q} - 1) \cdot (1 - q)^{-1} \tag{2.51}$$

and

$$S_q(r) = (N^{1-q} - 1) \cdot (1 - q)^{-1} \tag{2.52}$$

in the limit when $q = 1$, relation (2.51) or (2.52) leads to the Boltzmann-Shannon expression (Gell-Mann & Tsallis, 2004).

Below, a normalized entropy index is suggested, one in which the numerator stands for the calculated entropy of the system while the denominator displays the highest maximum entropy of the system owing to the equiprobability property:

$$S(\hat{a}_k) = -\left[1 - \sum_k \sum_m (p_{km})^q\right] / \left[k \cdot (M^{1-q} - 1)\right] \tag{2.53}$$

with k varying from 1 to K (number of parameters of the system) and m belonging to M (number of support space points), with $M > 2$. $S(\hat{a}_k)$ then reporting the accuracy on estimated parameters. Equation (2.54) reflects the non-additivity property of Tsallis entropy for two (probably) independent systems; the first, parameter probability distribution, and the second, error disturbance probability distribution (plausibly with quasi-Gaussian properties):

$$S(\hat{\mathrm{P}}\mathrm{r}) = [\, S(\hat{p} + \hat{r})] = \left\{[S(\hat{p}) + S(\hat{r})] + (1 - \mathrm{q}) \cdot S(\hat{p}) \cdot S(\hat{r})\right\} \tag{2.54}$$

where:

$$S(\hat{p}) = -\left[1 - \sum_k \sum_m (p_{km})^q\right] / \left[k \cdot (M^{1-q} - 1)\right]$$

and

$$S(\hat{r}) = -\left[\left(1 - \sum_n \sum_j r^q\right)\right] / \left[k \cdot N \cdot (J^{1-q} - 1)\right]$$

$S(\hat{\mathrm{P}}\mathrm{r})$ is then the sum of normalized entropy related to parameters of the model $S(\hat{p})$ and to the disturbance term $S(\hat{r})$. Likewise, the latter value $S(\hat{r})$ is derived for all observations n, with J the number of data points on the support space of estimated probabilities r related to the error term.

The values of these normalized entropy indexes $S(\hat{a}_k)$, $S(\hat{\mathrm{P}}\mathrm{r})$ vary between zero and one. Their values, near to one, indicate a poor informative variable while lower values are an indication of better informative parameter estimate $\hat{a}_k$ about the model.

The next part of the book will present in detail national accounts tables used for building or forecasting macroeconomic models. The statistical theory will be implemented particularly in the case of the inverse problem, while keeping in line with this work's objective.

2.7 An Application Example: a Maximum Tsallis Entropy Econometrics Model for Labour Demand

This example presents, through Monte Carlo simulations, a model for labour demand adjustment for the Polish private sector. It constitutes an extension of an initial model presented by Bwanakare (2010) for the labour demand adjustment by the private sector of Subcarpathian province in Poland. The model aims at displaying short-run and long-run relationships between labour demand determinants through an error self-correct process. Due to the relatively short period of the sample (fourteen annual data points) and the autoregressive nature of the model, we may have to deal with limited possibilities of statistical inference in the absence of convergence properties or, in the worst case, an inverse ill-behaved problem. Thus, traditional methods of parameter estimation may fail to be effective. We then propose to apply the generalized maximum Tsallis entropy econometric approach—as an extension of Jaynes-Shannon-Gibbs Information theoretic entropy formalism, already applied in econometrics (Golan, Judge & Miller, 1996). Due to an annual data frequency of the sample, the approach proves to be applicable in the case of classical econometrics when a small, lower frequency data sample is available. Such a small data sample should display tail queue Gaussian distribution. Through this application, Monte Carlo experiment outputs seem to confirm the reliability of the Tsallis entropy econometrics approach, which in this particular case performs as well as the generalized least square technique.

2.7.1 Theoretical Expectation Model

In the short run, managers decide on the number of employees to be hired (or dismissed) in accordance with the expected long-run optimal level of production. However, because of institutional or economic reasons, that optimal number is not hired (or fired) at once. First, uncertainty remains a predominant characteristic of business. For this reason, employers naturally prefer a moderate and progressive adjustment of recruited workers to the targeted optimal level. Recruitment in some economic sectors could be time-consuming as well, especially when searching for good specialists. Second, relatively well organized trade unions could prevent employers from abrupt, large-scale layoffs, or the cost of dismissing a worker may become high, depending on prevailing labour laws at a given period. In both cases, the process of shock correction will be more or less long, depending on its origin and magnitude.

Under classical assumptions of constant returns to scale, *ex ante* and *ex post* complementarities of factors, and long-run constant rate of labour productivity, the desired level of labour demand L_t^* is a function of the output Y_t and the technical progress t[20]:

$$L_t^* = \alpha.\exp(-\beta.t).Y_t \tag{2.55}$$

Assuming that labour demand adjusts to its targeted level by an error correction model:

$$\log(L_t / L_{t-1}) = \lambda.\log(L^*_t / L^*_{t-1}) + \mu.\ \log(L^*_{t-1}/L_{t-1}), \tag{2.56}$$

combining (2.55) and (2.56) leads to:

$$\log(L_t / L_{t-1}) = \lambda.\log(Y_t / Y_{t-1}) + \mu.\ \log(Y_{t-1}/L_{t-1}) + \mu.\ \beta.t + \alpha_o \tag{2.57}$$

The parameter λ is the impact of output on labour demand, and then a short-run elasticity of labour demand with respect to output Y_t, μ being the error correction parameter. Since a relation $-1 \le \mu \le 0$ should prevail, the equilibrium error is only partly adjusted at each period. In other words, this parameter synthesizes employers' determinants of labour demand adjustment once a shock in sales for the coming period is expected.

2.7.2 A Generalized Non-Extensive Entropy Econometric Model

2.7.2.1 General Model

Presently we are interested in the estimation of parameters of a Podkarpacki labour demand model, applying a generalized non-extensive entropy econometric approach. Following Golan, Judge & Miller (1996) and Bwanakare (2014a, 2014b), we reparametrize, in the first step, the generalized linear model before fitting it to Equation (2.48). This step allows for including in moment equations-restrictions the same probability variables as those optimized in the criterion function.

To reparametrize the model, we follow each equation in (2.45–2.46) where each β_k (k = 1,…,K) is treated as a discrete, random variable with compact support and $2 < M < \infty$ possible outcomes. Next, for the estimation of the model, we maximize the entropy criterion function in (2.47) under moment and normality condition restrictions presented in (2.48–2.50). For confidence area analysis, we need to apply Equations (2.53–2.54).

20 This is a simplification stipulating that technical progress is a linear function of time.

With the purpose of improving estimated parameter quality, one can add additional *a priori* restrictions to (2.48–2.50) as follows:

$$e = Y - Y = \hat{Y} - XV\hat{p} = 0. \qquad (2.58)$$

Then we constrain the error term e to sum up to zero[21] which provides an additional quality of requiring an unbiased parameter estimator.

The property of efficiency mainly depends upon the informative quality of both the prior (support space) and the model (econometric equation). When it is poor, the values of the estimated $\hat{p}_i$ from the model tends to be equal for all p_i, i.e., the case of a uniform distribution.

According to economic theory, we constrain elasticity parameters within a point support space of *zero* and *one*. As known (e.g., Golan, Judge & Miller 1996), sharper support area points of a parameter act as increasing quality of the "*a priori*" information. Furthermore, this allows computations of this nonlinear model to promptly converge to a global optimum solution. This is explained as follows:

$$0 \le \lambda = V\hat{p} \le 1 \qquad (2.59)$$

Likewise, we may add additional economic restrictions to the model (2.57) parameters; this leads to the following formulations:

$$-1 \le \mu = V\hat{p} \le 0 \qquad (2.60)$$

$$-\infty \le \beta = V\hat{p} \le 0 \qquad (2.61)$$

2.7.3 Estimated Confidence Area Of Parameters

In classical econometrics, we usually combine the variance of random model error with the co-linearity level of explicative variables to determine the standard error of estimated parameters and to infer their confidence area while assuming a normal distribution law of random errors. This is particularly true in the case of the Least Squares approach for a linear model.

In entropy econometrics, the approach is very different. We use the normalized entropy $S(\hat{a}_{ij})$ (Equation 2.53) as an equivalent of the estimate standard error measure in classical linear model econometrics. Likewise, the equivalent of the coefficient of determination R^2 is a $S(\widehat{\Pr})$ (Equation 2.54). Following Golan et al. (1996a, 1996b, 1996c, 2002) and Soofi (1992, 1994), in the case of maximum entropy formulation, the maximum level of entropy-uncertainty results when the information-moment

21 Note that our model has a constant term, suggesting that the economic initial condition may impact the optimal solution.

constraints are not enforced. Furthermore, this leads to uniform distribution of probabilities over the k states of the system. As we add each piece of informative data in the form of a constraint, a departure results from the uniform distribution, which explains an uncertainty reduction. Thus, the value of $S(\hat{P})$ reflects a global departure from the maximum uncertainty for the whole model. A similar measure, $1 - S(\hat{P})$, called *the information index*, explains the level of informative content of the model. For theoretical details, we refer the reader to the formulations presented above in Equations (2.51–2.54) or, e.g., in Golan et al. (1996).

2.7.4 Data and Model Outputs

In this section, the output parameters of Tsallis entropy, Shannon entropy, and least squares econometric models are presented. Next, the obtained results will be compared to those from a Monte Carlo simulation using the same data.

Data used in the model (Equation 2.57) come from the Polish Office of Statistics (GUS) and concern the period 1997–2010. Parameters of the model have been computed with the GAMS (General algebraic modelling system) code with the incorporated solver *PATHNLP*. We have noticed, through different simulations, that the Shannon-Gibbs entropy model seems more sensitive to initial conditions (support space of parameters in particular) than Tsallis entropy. This is a useful property, particularly when an economic theory does not exist to prompt us as to the starting parameters with which to begin. Parameter estimation by the robust standard errors least squares (LS) approach has been carried out, using freeware *Gretl* software *(http://gretl.sourceforge.net/)*. Thus, the HAC estimator is used for heteroscedasticity and autocorrelation correction.

a) Parameter outputs of Tsallis entropy model:

Dependent variable: $\log\left(L_t / L_{t-1}\right)$

Exogenous variables	$\log\left(y_t / y_{t-1}\right)$	$\log\left(Y_{t-1} / L_{t-1}\right)$	T	a0
Estimates $\hat{a}_j$	0.710	0.010	-0.020	-0.266
Precision error $S(\hat{a}_k)$ on estimated parameters	0.135	0.250	0.250	0.236

Information Index $I[S(\hat{P}r)] = 1 - S(\hat{P}r) = 0.852$
Tsallis -q parameter (for a weight αi=15%) = 2.091

Throughout many conducted experiments, we have observed the coefficient $S(\hat{\text{P}}\text{r})$ to be very sensitive to weighting parameters α in the objective function. Tsallis-q value being itself influenced by the above weights, its values closer or higher to 5/3 correspond to meaningless information index coefficients for which $S(\hat{\text{P}}\text{r})$ vanishes to zero. In empirical research, the Tsallis-q coefficient may take much higher values as a consequence of model linearity attributes or in the case when the sample is small. In the present case, we have noticed a high sensitivity of this Tsallis-q parameter on the change of the weight α_i in the criterion function. The higher the weight $\alpha_{i,}$ the higher the value of the Tsallis-q parameter. We have retained the value of this weight for which $I[S(\hat{\text{P}}\text{r})]$ is the highest.

b) Parameter outputs of Shannon-Gibbs entropy model:

Dependent variable: $\log\left(L_t / L_{t-1}\right)$

Exogenous variables	$\log\left(y_t / y_{t-1}\right)$	$\log\left(Y_{t-1} / L_{t-1}\right)$	T	a0
Estimates $\hat{a}_j$	0.709	0.010	−0.020	−0.263
Precision error $S(\hat{a}_k)$ on estimated parameters	0.297	0.518	0.518	0.421
Information Index $I[S(\hat{\text{P}}\text{r})] = 1 - S(\hat{\text{P}}\text{r}) = 0.829$				

c) Robust standard errors LS estimation:

Dependent variable: $\log\left(L_t / L_{t-1}\right)$

Exogenous variables	$\log\left(y_t / y_{t-1}\right)$	$\log\left(Y_{t-1} / L_{t-1}\right)$	T	a0
Estimates $\hat{a}_j$	0.709963	0.010417	−0.02031	−0.26578
P-values	2.73e−011***	4.34e−06	2.89e−06***	0.0302**
Corrected R2 = 0.79 DW = 1.4832				

Three parameters are different from zero at 1%, and one on the variable a_0 significant at 10%. The above precision on the estimated parameters from such a small data sample of an autoregressive model suggests the presence of co-integrating—at the same order—variables L_t and Y_t. Such a particular situation leads to super-consistency of estimated parameters.

Table 3: Monte Carlo simulation outputs

	Poland Labor demand Model and simulation outputs						
		NEE/GLS	#5000	#10000	#15000	#20000	#25000
Estimates	y_t / y_{t-1}	0.709963106	0.7092748	0.70445994	0.7126441	0.709756	0.708636
	Y_{t-1} / L_{t-1}	0.010416764	0.0105625	0.01052379	0.0103705	0.010461	0.010458
	T	-0.020312463	-0.020624	-0.0205689	-0.020213	0.020406	-0.02038
	C_0	-0.26577817	-0.272354	-0.2652226	-0.266263	-0.26784	-0.26663
Standard errors	y_t / y_{t-1}	0.176056192	0.3303391	0.33502592	0.3377129	0.331089	0.335143
	Y_{t-1} / L_{t-1}	0.003753087	0.0070234	0.00704595	0.0071221	0.007143	0.007144
	T	0.007938838	0.0148845	0.01489532	0.0150799	0.015074	0.015107
	C_0	0.222643946	0.4216688	0.42344287	0.4260381	0.420446	0.426712

Source: own elaboration.

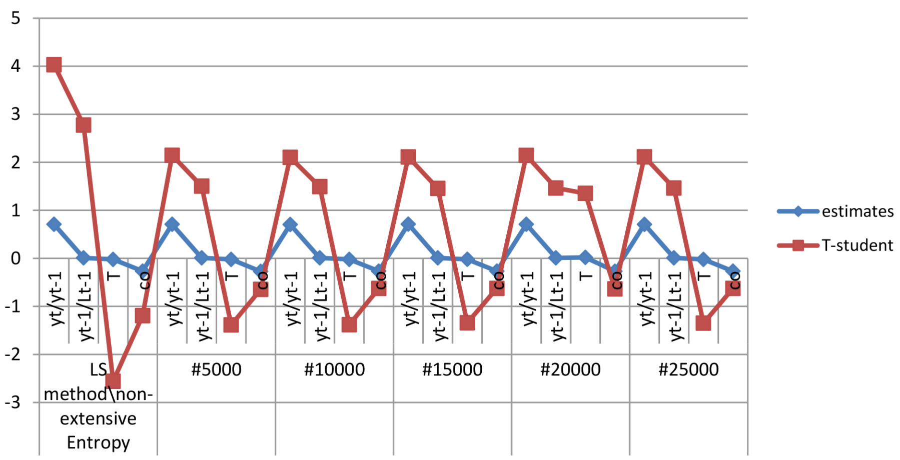

Figure 1: Monte Carlo model estimates and T-student from simulations: initial model, #5000, #10000, #15000, #20000, #25000.

For comparative purposes, Table 3 presents outputs from Monte Carlo experiments (computed with *Mathlab* 7.3.0 software).

The above outputs have been derived under the hypothesis of random normal law. The empirical standard error initially computed for random value generation is 0.02035. This constitutes 50% of the observed endogenous variable standard error. We observe that Shannon-Tsallis entropy and the least squares outputs are similar and almost reflect Monte Carlo convergence outputs. The initial t-student related to the parameters on the variables y_t/y_{t-1} and y_{t-1}/L_{t-1} decrease when we carry out the #5000 simulation and remain practically unchanged up at the #25000 simulation experiment. Nevertheless, we observe that parameter estimates of the model remain unchanged irrespective of the number of the simulations.

To conclude, we note the accuracy in the similarity of outputs from the three models. This suggests that we are dealing with a convergent case of power law to Gaussian distribution. If the Tsallis-q parameter is too high, it cannot be interpreted in a model where its nonlinearity and the small sample size (in this case 14 observation years) should have a significant impact on the value of that parameter (Grech & Pamula, 2013). The impact parameter is around 0.71. This is, on average, a 0.71% growth of labour demand when gross profits shift up to 1%. As it has been indicated, these outputs are related to a period (1997–2010) during which Poland was undergoing structural, post-communism reforms. As such, their interpretation should be done carefully. As far as exogenous technical progress is concerned, we observe a negative sign on the value of the estimated parameter β on the symptomatic variable t, which indicates an expected adverse impact of technical progress on labour demand.

Annex A

The solution for the above constrained equation is obtained by forming the Lagrange function:

$$L = -\sum_{k=1}^{K} p_k \ln p_k + \sum_{t=1}^{T} \lambda_t \left[y_t - \sum_{k=1}^{K} p_k f_t(x_k) \right] + \mu \left(1 - \sum_{k=1}^{K} p_k \right) \tag{2.20}$$

After defining the first order conditions, the solution to this (K + T + 1) equations and parameters is:

$$\hat{p}_k = \exp\left(-\sum_{t=1}^{T} \hat{\lambda}_t f_t(x_k) - 1 - \hat{\mu} \right), k = 1,2,\ldots K,$$

$$\sum_{k=1}^{K} \exp\left(-\sum_{t=1}^{T} \hat{\lambda}_t f_t(x_k) - 1 - \hat{\mu} \right) f_t(x_k) = y_t, t = 1,2,..,T,$$

$$\sum_{k=1}^{K}\exp\left(-\sum_{t=1}^{T}\hat{\lambda}_t f_t(x_k)-1-\hat{\mu}\right)=1,$$

and, as a formal solution, we obtain:

$$\hat{p}_k=\frac{1}{\Omega(\hat{\lambda}_1,\hat{\lambda}_2,\ldots\hat{\lambda}_T)}\exp\left(-\sum_{t=1}^{T}\hat{\lambda}_t f_t(x_k)\right),$$

where

$$\Omega(\hat{\lambda})=\sum_{k=1}^{K}\exp\left(-\sum_{t=1}^{T}\hat{\lambda}_t f_t(x_k)\right)$$

is a normalization factor.

It is easy to prove the uniqueness of the primal ME solution. In fact, given the first order conditions, elements of the Hessian matrix are as follows:

$$\frac{\partial^2 L}{\partial p_k{}^2}=\frac{-1'\exp(f(x_k)'\lambda)}{\exp(f(x_k)'\lambda)}=-\frac{1}{p_k} \text{ for the diagonal elements,}$$

and

$$\frac{\partial^2 L}{\partial p_k \partial p_j}=0 \text{ for the off-diagonal elements.}$$

Thus, the Hessian is negative defined and sufficient condition for a unique global maximum is fulfilled. Furthermore,

$$y_t=\left(\frac{\partial}{\partial \lambda_t}\right)\ln\Omega, 1\le t\le T$$

One may observe that the value of the entropy H is a function of the given data:

$$H=\ln\Omega(\hat{\lambda})+\sum_{t}^{T}\hat{\lambda}_t y_t\,, \tag{2.21}$$

Annex B: Independence of Events Within q-Generalized Kullback-Leibler Relative Entropy

Let us consider (Tsallis, 2009) the problem of independence of random variables in the case of two-dimensional random variable (x, y), and its corresponding distribution function $p(x, y)$ with $\int dxdyp(x, y) = 1$

As expected, the marginal distribution functions are then given by

$$h_1(x) \equiv \int dyp(x, y)$$

and

$h_2(y) \equiv \int dxp(x, y)$.

The discrimination criterion for independence concerns the comparison of $p(x, y)$ with $p_o(x, y) \equiv h_1(x)h_2(y)$. Once again, the one-dimensional random variables x and y are independent if and only if $p(x, y) = p_o(x, y)$. Therefore, the criterion becomes:

$$\int dxd(y)\frac{\left[\frac{p(x,y)}{h_1(x)h_2(y)}\right]^{q-1}-1}{q-1} \geq 0, \text{ for } q \succ 0.$$

When $q \to 1$, we then recover the usual discrimination criterion, i.e.:

$$\int dx\, dy\, p(x, y) \ln p(x, y) - \int dx\, h_1(x) \ln h_1(x) - \int dy\, h_2(y) \ln h_2(y) \geq 0.$$

An interesting case is if $q \to 2$, then we have:

$$\int dxd(y)\left[\frac{p(x,y)}{h_1(x)h_2(y)}\right]^2 \geq 1,$$

The value of this quantity, useful in economics, may give a sign of independence between x and y, when it vanishes.

Finally, if we generalize to the case of many variables, the Kullback-Leibler-Tsallis index of information becomes:

$$Iq\ (p(x1, x2, \ldots, xd), p_o(x1, x2, \ldots, xd)) \geq 0\ (\text{for } q \geq 0)$$

or its symmetrized version:

$$\frac{1}{2}\ [I_q\ (p(x_1, x_2, \ldots, x_d), p_o\ (x_1, x_2, \ldots, x_d) + I_q p_o\ (x_1, x_2, \ldots, x_d)\ p(x_1, x_2, \ldots, x_d))] \geq 0\ (q \geq 0)$$

When equality holds for these two above relations, this means that all elements $x_1, x_2, \ldots, x_d$ are independent among them (almost everywhere).

Bibliography – Part II

Abe, S. (2009). Generalized molecular chaos hypothesis and the H theorem: Problem of constraints and amendment of nonextensive statistical mechanics. *Phys. Rev.* E 79 (4), 041116.

Abe, S., & Bagci, B.G. (2005). Necessity of q-expectation value in nonextensive statistical mechanics. *Phys. Rev.* E 71(1), 016139.

Braaksma, B., & Zeelenberg, K. (2015). "Re-make/Re-model": Should big data change the modelling paradigm in official statistics? *Statistical Journal of the IAOS* 3:193–202.

Bwanakare, S. (2010). Non-Extensive Entropy Econometric Model (NEE): the Case of Labour Demand in the Podkarpackie Province. *Acta Physica Polonica A*, 117(4):647–651.
Bwanakare, S. (2014a). Econometric Balancing of a Social Accounting Matrix Under a Power-law Hypothesis. *Przegląd Statystyczny* 61(3):263–281.
Bwanakare, S. (2014b). Non-Extensive Entropy Econometrics: New Statistical Features of Constant Elasticity of Substitution-Related Models. *Entropy* 16(5):2713–2728.
Bwanakare, S. (2015). Greenhouse Emission Forecast as an Inverse Stochastic Problem: A Cross-Entropy Econometrics Approach. *Acta Physica Polonica A* 127(3A):13–20.
Carter, T. (2011). *An introduction to information theory and entropy*. Complex Systems Summer School. Retrieved from: http://astarte.csustan.edu/~tom/SFI-CSSS
Drożdż, S., & Kwapień, J. (2012). Physical approach to complex systems. *Physics Reports* 515(3–4):115–226.
Engle, R.F., & Granger, C.W.J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2):251–276.
Gell-Mann, M., & Tsallis, C. (Eds.). (2004). *Nonextensive Entropy, Interdisciplinary Applications*. USA: Oxford University Press.
Golan, A., Karp, S.L., & Perloff, J.M. (1996a). *Estimating a Mixed Strategy Employing Maximum Entropy*. eScholarship, California. Retrieved from http://escholarship.org/uc/item/5rx8z14p
Golan, A., Judge, G.G., & Miller, D. (1996b). *Maximum Entropy Econometrics: Robust Estimation with Limited Data.* Chichester, England: Wiley.
Golan, A., Judge, G.G., & Perloff, J.M. (1996c). Estimating the Size Distribution of Firms Using Government Summary Statistics. *The Journal of Industrial Economics*, 3(1):69–80.
Golan, A., & Perloff, J.M. (2002). Comparison of Maximum Entropy and Higher-Order Entropy Estimators. *Journal of Econometrics,* 107(1):195–211.
Good, I.J. (1963). Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables. *The Annals of Mathematical Statistics* 34(3):911–934.
Grech, D., & Pamula, G. (2013). On the multifractal effects generated by monofractal signals. *Physica A,* 392(23):5845–5864.
Green, W.H. (2003). *Econometric Analysis* (5th Ed.). NY, USA: Prentice Hall.
Halmos, P.R., & Savage, L.J. (1949). Application of the Radon-Nikodym Theorem to the Theory of Sufficient Statistics. *The Annals of Mathematical Statistics,* 20(2):225–241.
Hartley, R.V.L. (1928). Transmission of Information. *Bell Labs Technical Journal*, 7(3):535–563.
Jaynes, E.T. (1994). *Probability Theory: The Logic Of Science.* Washington DC, USA: Washington University.
Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems. *Proceedings of the Royal Society A,*186(1007):453–461. JSTOR 97883.
Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.,* 82(D):35–45.
Khinchin, I. (1957). *Mathematical Foundations of Information Theory.* NY: Dover.
Kullback, S. (1959). *Information Theory and Statistics,*. NY: John Wiley and Sons.
Landsberg, T.P., & Vedral, V. (1998). Distributions and channel capacities in generalized statistical mechanics. *Physics Letters A*, 247(3):211–217.
Loeve, M. (1955). *Probability Theory: Foundations, Random Sequences,*. NY: Van Nostrand.
Papoulis, A. (1991). *Probability, Random Variables and Stochastic Processes* (3rd Ed.). N.Y., USA: McGraw-Hill.
Plastino, A., & Venkatesan, R.C. (2011). Deformed Statistics Kullback-Leibler Divergence Minimization within a Scaled Bregman Framework. *Physics Letters A,* 375(48):4237–4243.
Pukelsheim, F. (1994). The Three Sigma Rule. *The American Statistician,* 48:88–91.
Rajagopal, A.K., & Abe, S. (2000). Justification of Power Law Canonical Distributions Based on the Generalized Central Limit Theorem. *Europhysics Letters,* 52(610).

Rényi, A. (1961). On measures of information and entropy. *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob.*, Vol. 1:547–561. (Univ. of Calif. Press, 1961).

Savage, L.J. (1954). *The Foundations of Statistics.* NY, USA: Wiley.

Shannon, C.E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal,* 27: 379–423 and 623–656.

Soofi, E.S. (1992). A Generalizable Formulation of Conditional Logit with Diagnostics. *Journal of the American Statistical Association,* 87:812–816.

Soofi, E.S. (1994). Capturing the Intangible Concept of Information. *Journal of the American Statistical Association,* 89:1243–1254.

Theil, H. (1967). *Economics and Information Theory.* Chicago, USA: Rand McNally and Company.

Tikhonov, A.N., & Arsenin, V. I. (1977). *Solutions of Ill-conditioned Problems.* New Jersey: John Wiley & Sons.

Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. Berlin: Springer.

Part III: **Updating and Forecasting Input-Output Transaction Matrices**

1 Introduction

The second part of this book was concerned with showing statistical theory of information-based approaches as a basis for solving the ill-behaved inverse problem. This third part deals with the applications of the statistical theory of information in macroeconomics. It introduces *ad hoc* macroeconomic theory before building and estimating different models in the context of an inverse problem. A system of national accounts with particular emphasis on input-output tables reflecting complex interactions between economic activities, product markets, factors of production, and the behaviour of different economic agents is presented. A section is devoted to input-output multipliers. Numerical examples are provided using the RAS approach as well as the non-extensive entropy econometrics procedure. Limited extension of these tables will allow us to treat certain themes of the natural environment through a theoretical model. Then, tables such as input-output tables are treated in detail in the context of modelling or forecasting when prior information is insufficient and the matrix is ill-behaved. Faced with this problem, maximum entropy formalism, in particular, Tsallis non-extensive entropy, will be employed.

2 The System of National Accounts

The System of National Accounts (SNA) constitutes the internationally agreed set of standardized conventions on how to compile, in a coherent and consistent way, an integrated set of macroeconomic accounts to measure economic activity (e.g., Eurostat, 2002). Such procedures require a set of internationally agreed-upon concepts, definitions, classifications, and accounting rules. In addition, the SNA provides an overview of economic processes, recording how production is distributed among consumers, businesses, government, and foreign nations. It shows how income originating in production, modified by taxes and transfers, flows to these groups and how they allocate these flows to consumption, saving, and investment. Consequently, the national accounts are building blocks of macroeconomic statistics, forming a basis for economic analysis and policy formulation.

The SNA is intended for use by all countries, having been designed to accommodate the needs of countries at different stages of economic development. It also provides an overarching framework for standards in other domains of economic statistics, facilitating the integration of these statistical systems to achieve consistency with the national accounts. However, the complexity of the interrelations that emerge makes the understanding of underlying macroeconomic rules difficult.

In 1947, Stone (1955, 1981), then head of the League of Nations Committee of Statistical Experts, submitted for the first time a report from the Subcommittee on National Income Statistics that would constitute the origins of the SNA. During the same year, the United Nations Statistical Commission (UNSC) promoted the evident need for international statistical standards for statistical comparisons in support of a large array of policy implementations.

This led in 1953 to the publication of the SNA under the auspices of the UNSC. It consisted of a set of six standard accounts and a set of twelve standard tables presenting details and alternative classifications of flows in the economy. Successive modifications and extensions were implemented in 1960, 1964, 1968, 1993, and 2008. Extensions made in 1968 deserve more attention in the context of the present monograph. Input-output accounts and balance sheets were added to the framework of the SNA. Attention was focused on constant price derivation and a comprehensive effort was deployed to bring the SNA and the Material Product System (MPS) closer together. The archetype of the modern SNA was implemented in 1993 (Beutel & De March, 1998). In fact, the 1993 SNA represented a major advance in national accounting and embodied the result of harmonizing the SNA and other international statistical standards more completely than in previous versions. Since the first United Nations Conference on the Human Environment held in Stockholm in mid-1972, increasing needs to incorporate environmental aspects into the SNA started to be fulfilled in 1993 when a system of integrated environmental and economic accounting (SEEA) was introduced by Caticha & Giffin (UN, 1993; USA, 2007). The 2008 SNA update addresses issues brought about by changes in the economic environment, advances in methodological research, and the needs of users.

3 The Input-Output (IO) Table and its Main Application

Leontief (1941, 1966), the 1973 Economics Noble laureate and the father of the I-O table approach to economic analysis, began his first book about I-O analysis with the following words:

> This modest volume describes an attempt to apply the economic theory of general equilibrium—or better, general interdependence—to an empirical study of interrelations among the different parts of a national economy as revealed through covariations of prices, outputs, investments, and incomes.

Leontief tried to apply neo-classical (Walras) general equilibrium to practical economic life. This suggests that subsequent analyses based on I-O tables or their extensions could have economic interpretations within the Walrasian framework apart from a few particular cases—e.g., those that consider the environment—violating Pareto optimum conditions.

The objective of this chapter is to present a consistent methodology of updating, forecasting, and economic modelling on the basis of I-O tables—for which underlying matrices are ill-behaved or data are not reliable. The proposed maximum entropy methodology can dynamically assess I-O multipliers and update and forecast I-O table information by combining the generalized maximum entropy principle and macroeconomic theory. The procedure remains in line with multiplier-accelerator analysis, assuming that induced investment is a function of expected growth. The only required condition to apply the proposed techniques is the availability of statistical information on final demand or value-added accounts which allow for updating under some constraining information (macroeconomic or not) obtained earlier, according to the traditional approach I-O table.

In the following sections, classical structure of an I-O table will be reviewed. The next step will describe I-O multipliers and their usage before trying to solve the more complex aspects of their estimation. Next, the proposed methodology for updating and forecasting an ill-behaved IO table will be described and the model presented.

3.1 The I-O Table and Underlying Coefficients

The I-O method (Tomaszewicz 1992, 2005) represents a quantitative research approach that helps to understand how economic global product is created and shared with particular reference to connections within different sectors of production at this intermediate stage of product processing (Almon & Clopper, 2000; Avonds & Luc, 2007; Robinson, Cattaneo & El-Said, 2001).

Use of this method of analysis is built upon I-O tables, the construction of which assumes the existence of constant returns to scale in the production process and the existence of general Walrasian equilibrium in the overall economy. In fact, as suggested earlier, the intention of Wassily Leontief was to apply general equilibrium theory previously proposed by Warlas (Leontief, 1970, 1986). In the context of the methodology of this work, let us first present the familiar I-O table structure, known for many decades, with only a slight modification that allows it to retain its square form. The fundamental concept of the I-O model is the concept of a direct technical financial coefficient:

$$a_{ij} = \frac{x_{ij}}{X_j} \qquad \text{for } i,j = 1,\ldots,n. \tag{3.1}$$

Matrix $A = (a_{ij})$ is called a matrix of coefficients of inputs. It expresses the proportion of the value of product from sector *i* to be involved (sold) in the sector *j* for production of one unit value (e.g., 1 euro). The elements of this table are also called technical coefficients when expressed in quantitative flows between industries.

Table 4: Simplified inputs-outputs table structure.

Sector	Flows					Final demand	Total
1	x_{11}	x_{12}	x_{13}	…	x_{ln}	Y_{1f}	X_1
2	x_{21}	x_{22}	x_{23}	…	x_{2n}	Y_{2f}	X_2
…	…	…	…	…	…	…	….
N	x_{n1}	x_{n2}	x_{n2}	…	x_{nn}	Y_{nf}	X_n
	D_1	D_2	D_3	…	D_n		D
Total	X_1	X_2	X_3	…	X_n	Y_f	

Source: own elaboration.

Indications:
X_j is the value of the global product of *j*- branch, $j = 1, \ldots, n$.
x_{ij} is the flow from the branches *i* to *j*, i.e., the value of the product manufactured in branch i-th and consumed by a branch of the j-th, $i, j = 1, \ldots, n$.
Y_{if} is the value of the final demand, $i = 1, \ldots, n$ and f =1, …*F.* Index *f* represents final demand composition, such as households, investment, exports, etc. The number *F* depends on the degree of table aggregation.
D_j is the value added from branches of the *j*-th, $j = 1, \ldots, n$.

Note that the above table can be split into four main parts:

The first part (upper-left part of the table) is composed of sub-matrix *A* showing the structure of intermediary demand of products "*i*" by sector "*j*".

In the second part (top right), we have the final demand (Y_{if}). Its traditional elements are: household consumption, government consumption, investment and stocks and the export sector.

The third part (lower left) shows the revenue created in the modes of production, i.e., the value added (D_j), i.e., remuneration of labour and gross profit of capital in production sectors. If in the second part the "export" sector is explicitly shown, then in the third part one must, consequently, show the import sector.

Part four of the table refers to the secondary division of generated revenues. In the case of the I-O matrix it remains empty. Only construction of the social accounts matrix allows for completing this information.

Here, one should recall macroeconomic balance between final demand and value added, i.e.:

$$\sum_j D_j = \sum_i Y_i .$$

A further important equation is the definition of the Leontief model:

$$(I - A)X = Y \tag{3.2a}$$

or

$$(I - A)^{-1} Y = X \tag{3.2b}$$

where:
$X = [X_j]$ n-dimensional relationships to the column vector of global product
$Y = [Y_i]$ n-dimensional relationships to the row vector of the final product, and I is the identity matrix. This relationship between the final product Y, the global product X, and cost structure matrix A is useful for the calculation of one of these elements when information about the other items is available. In the literature, this is known as a forecast of the first type (3.2a) or the second type (3.2b).

3.2 Input-Output Multipliers

3.2.1 The principal models

Multiplier coefficients play a central role in economic analysis (Leontieff, 1970). They make it possible to measure the impact of the exogenous variable or a shock on the whole system in which elements are interconnected. On these grounds, systemic models like these based on I-O tables—or their extension, or on computable general equilibrium models—have proved decisively superior to the classical *ceteris paribus* approach, using the classical econometric models.

The above defined relations (3.2a) and (3.2b) are very useful in empirical research. (3.2b) explains, in (input) multiplier terms, responses of producing sectors to a one unit value increase in sectorial demand. For example, if the government spends one additional euro for buying a product from a given sector, what will the (backward) production impact be on the whole economy? The response is given by the level of multiplier coefficients along the sector column under consideration and total impact is given by its multiplier sum. The term $(I - A)^{-1}$ in relation (3.2b) represents the multiplier. This formula can be decomposed as follows:

$$(I - A)^{-1} = \sum_{i=0}^{i=\infty} A^i \text{ with } A^0 = \mathrm{I} \text{ and } A^i = A^{i-1} A \tag{3.2c}$$

Such a multiplier displays three impacts:

– initial impact equal to one,
– direct impact equal to A, and
– indirect impact summing up to $A^2 + A^3 + \ldots + A^k + \ldots + A^n$. Note that this geometric progression is quickly convergent; after a few steps, the last term vanishes to zero.

If we try to make a thorough analysis regarding the probabilistic nature of the multiplier, then we can rewrite (3.2c) as follows:

$$e^A i! \cong \sum_{i=0}^{\infty} A^i = (I - A)^{-1}, \tag{3.2d}$$

after having used the Taylor development. Thus, due to (3.2d), economic multiplier structure displays the exponential family of laws, which may constitute—as indicated in the first part of this book—a transitional law between power law and Gaussian law when power law-related higher frequency data are progressively aggregated towards the low frequency series. The purpose of this short discussion is to draw attention to the use of formula (3.2c). Not only should the frequency level of data have an impact on results, but also the functional relation of matrix A could modify the convergence transition path between the three probabilistic laws above.

Equation (3.2b) is generalized by the following, one of the most important relations of I-O modelling theory, referred to as *the central model of input-output*:

$$Z = B\,(I - A)^{-1}\,Y \tag{3.3}$$

B = matrix of input coefficients for specific variables in economic analysis (intermediates, labour, capital, energy, emissions, etc.)
I = Identity matrix
A = matrix of technical financial coefficients $[a_{ij}]$
Y = diagonal matrix of final demand
Z = matrix with results for direct and indirect requirements (intermediates, labour, capital, energy, emissions, etc.)

Three of the most frequently used types of multipliers in I-O analysis are those that estimate the effects of the exogenous changes of final demand (consumption, investment, exports) on:
a) outputs of the sectors in the economy,
b) value added and income earned by the households, and
c) employment that is expected to be generated by the new activity levels.

However, due to interesting potential applications of this theory, it is worthwhile to be more complete about I-O multipliers. In empirical research, the I-O models used are based on input coefficients. Nevertheless, there is also a family of I-O models that are based on output coefficients. These models are sometimes called Gosh-models (Gosh, 1964). These models can be used to study price and cost effects or forward linkages of industries. Input coefficients reflect production functions or cost structures of activities. In contrast, output coefficients are distribution parameters for products reflecting market shares.

Presented below are only the four basic I-O models with input and output multipliers. The four I-O models have a dual character with an underlying symmetry. Each I-O model with input coefficients has a complement with output coefficients.

$d_j = D_j / X_j$, i.e., input coefficient for value added. (3.4)

Input coefficients for intermediates (a_{ij}) (3.1) reflect the requirements for the use of product *i* in industry *j* for one unit of output of industry *j*. The capital and labour requirements are defined in the same way.

$b_{ij} = X_{ij} / X_i$, output coefficients for products (3.5)

$y_i = Y_i / X_i$, output coefficients for final demand (3.6)

Output coefficients for intermediates (b_{ij}) identify the share of deliveries of sector *i* for sector *j*, (x_{ij}) in the total output of sector *i*.

Model 1: quantity model with input coefficients
This model has been already defined in (3.2a) and (3.2b).

Model 2: price model with input coefficients

$A'p + d = p$ (3.7)
$(I - A')p = d$ (3.8)
$p = (I - A')^{-1}d$ (3.9)

A' = transposed matrix of input coefficients for intermediates with $A = [a_{ij}]$ for i, $j = 1,\ldots,m$.
I = identity matrix

x = column vector of unit product price indexes for sectors 1 to m. w = column vector of exogenous input coefficients for value added $w_1,\ldots w_m$.

Model 3: price model with output coefficients

$$Bp + y = p \tag{3.10}$$
$$(I - B)p = y \tag{3.11}$$
$$p = (I - B)^{-1}y \tag{3.12}$$

B = matrix of output coefficients for intermediates with $B = b_{ij}$ for $i, j = 1,\ldots,m$
I = identity matrix
p = column vector of unit product price indexes for sectors 1 to m
y = column vector of exogenous output coefficients for final demand by product, with $y_1,\ldots,y_m$.

Model 4: Quantity model with output coefficients

$$B'x + Z = X \tag{3.13}$$
$$(I - B')X = Z \tag{3.14}$$
$$x = (I - B')^{-1}Z \tag{3.15}$$

B' = Transposed matrix of output coefficients for intermediates with $B = b_{ij}$ for $i, j = 1,\ldots,m$
I = identity matrix
X = Column vector of product output for sectors 1 to m
Z = Column vector of exogenous value added by sector, with $Z_1, \ldots, Z_m$.

However, due to lack of a proper microeconomic foundation, I-O models with output coefficients are rarely used in empirical research. The question often put to empirical researchers concerns the extent to which I-O multipliers are stable in time and behave according to expectations. Naturally, we can ask the same—but on a different scale—concerning the stability of I-O coefficients. We do not pretend to have the answers to these questions, which remain beyond the scope of this book.

3.2.2 A Model of Recovering the Sectorial Greenhouse Gas Emission Structure

Starting from insufficient information, let us combine below the central model of I-O (3.3) with the statistical information theory approach to predict emission multipliers of an extended I- O table (Miller & Blair, 1985). As above, let vector X_t be the global product of a given economy, A_t a matrix of I-O coefficients, Y_t a diagonal matrix of final demand, and B_t a matrix of outputs emission. Let us now suppose that these pieces of information have been obtained for a given period zero t_0 and that this has made it possible to derive multipliers Z_0 from the equation explaining emission content from final demand.

Table 5: Example of matrix B_1 of ecological emissions.

Themes	Sector 1	Sector 2	…	Sector *n*	Themes total (period 1)
CO_2					E_1
CH_4					E_2
N_2O					E_3
HFC_S					E_4
PFC_S					E_5
SF_6S					E_6
Total global product	X_1	X_2	…	X_n	T

Next, we need to predict these multipliers for the next period one t_1 to eventually compare changes within economic sectors and/or environmental themes. As very often happens in the real world, the only information concerning each emission theme is the total E_i (where index i refers to each type of theme) estimated at the end of the period t_1, for example, a certain number of tons of CO_2 emitted during that period is equal to E_1 without, however, any knowledge concerning the industrial branch being the source of that emission. We must then estimate matrix B_1 (see Table 5) on the basis of the *a priori* information of the initial period 0 and from data on measured sector totals of each ecological theme in the current period 1.

Evidently, we are confronted with an inverse problem since there may exist infinite combinations of emission levels by theme, related to global product, the total of which could lead to E_i. Without additional imaginative assumptions, any classical approach could solve this class of problem. As we will see later, some techniques exist like the bi-proportional approach, known as the RAS method (Bacharach, 1971), which could offer a solution with sufficiently limited effects as it does not take into account additional information on the investigated stochastic model. The Bayesian approach could be used in this class of problem. It can be shown that the approach presented below may be seen as an enhanced Bayesian technique to incorporate the second law of thermodynamics, which is a natural principle of organization. The neuronal class of model could also be suggested. However, it is not based on any theory, its application is time-consuming, and its outputs are not always guaranteed.

Using minimum information and without additional assumption, we suggest solving the problem by applying minimum entropy formalism, according to (2.43 or 2.44)[22]. Shannon-Gibbs entropy has been applied with success for updat-

22 At the end of Part IV of the book, a theorem is presented proposing the power law properties of

ing and balancing matrices. However, on theoretical grounds, this assumes that entropy is a positive function of the number of possible states and is extensive (see 2.21), and it then neglects the possibility of inter-correlations among the states and their impact.

Let us then minimize the criterion function of the next non-extensive entropy model:

$$\text{Min}(p, p_0) \rightarrow I_q\left[p \| p_0\right] = \frac{1}{q-1}\sum_j p_j\left[(p_j)^{q-1} - (p_{0j})^{q-1}\right] - \sum_j (p_j - p_{0j})(p_{0j})^{q-1} \tag{3.16}$$

subject to

$$\sum\nolimits_{j=1}^{n}\sum_{h=1}^{M} p_{jh} = 1$$

$$\text{E}_\text{i} = \sum_{j=1}^{n}\sum_{h=1}^{M}\frac{p_{jh}{}^{q}}{\sum_{j=1..n} p_{jh}{}^{q}} Z_h X_j \tag{3.17}$$

Probabilities r_{ij} distribution of theme emission *i* on column *j* is defined on a discrete support space Z_h where *H* is the number of points minimally equal to two, $z = [z_1, z_2, \ldots, z_H]$. Support space is added since r_j are not probabilities and do not sum—by column—to unity. On a support space *Z*, parameters r_j are distributed with probabilities

$P_{jk} = [p_{jk1}, p_{jk2}, \ldots p_{jkH}]'$.

Thus, in matrix formulation, we have:

$E = ZP$,

with

$$Z = \begin{pmatrix} z' \ldots & & \\ & z'.. & \\ \ldots & & z' \end{pmatrix}$$

$$P = \begin{pmatrix} p_{11} & p_{21}\ldots & \ldots\ldots p_{H1} \\ p_{12}.. & p_{22}\ldots & ..p_{H2}. \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ P_{1H} & ..P_{2H} & ..\;p_{HH} \end{pmatrix}$$

---

an input-output or a SAM. For the time being, it is suggested that the existence of these properties be assumed and non-extensive entropy formalism be applied. As we already know, non-extensive entropy formalism generalizes Shannon-Kullback-Leibler, which means at least we cannot lose the advantages of that generalization.

Moments expression in (3.17), as already seen in Part II of this work, contains a term of probability coefficients referred to as *escort distribution*[23].

In that way we get quantity emission *eij*, being elements of matrix B, and on this basis, we can immediately derive the matrix of direct and indirect requirement emissions consecutive to one unit final demand Y_i according to the relation (3.3).

In this example, we have supposed the global product x_j and transactions matrix x_{ij} or I-O coefficients are known. At the end of this book, in Annex C (Tables 23 and 24), an instructional example is provided in which we are asked to recover total pollutants emission by industrial sector and by region on the basis of moment information. In the next section, we will remove these assumptions and suppose that the only information we have about the current I-O table is the final demand vector Y and *a priori* components, a previous I-O matrix.

3.3 Updating and Forecasting I-O Tables

3.3.1 Generalities

Methods for updating I-O tables (Snower, 1990; Toh, 1998) can be categorized into univariate, bivariate, econometric, and stochastic procedures (e.g., Miller & Blair, 1985:266–316). All methods attempt to solve the following problem: row and column sums of an I-O table should correspond to the exogenous projection, and negative inputs should be avoided.

The basic idea of univariate methods to update I-O tables is to correct the matrix of input coefficients row-wise with a diagonal matrix of correction factors. An example of a version of such a method is the Statistical Correction Method.

The Bivariate approach, in contrast to univariate methods, which work with corrections of rows only, corrects rows and columns of an I-O table at the same time. The well-known RAS approach (Stone, 1984; Toh, 1998) represents an example of such an approach. However, a simple RAS procedure will normally fail to produce an acceptable projection of the structural change of an I-O table when change in relative prices and change in technology are substantial. Nevertheless, the incorporation of other exogenous data in the modified RAS procedure will tend to improve the quality of the projection. Several variations of the RAS technique can be found, for example, in Miller and Blair (1985:276–313).

Stochastic procedures assume that many independent variables may influence changes of input coefficients. The changing coefficients do not follow homogenous

---

23 As discussed in (II.2.5.2), note that there is now an ongoing discussion of whether or not his form of constraint is appropriate in this case of Bergman relative entropy.

row and column multipliers, but rather the complex features of stochastic elements. The Lagrange method applied by the Central Statistical Office of the Netherlands (CBS) is an example of that method. The Least Squares Method can constitute another example of such an approach. It is worthwhile to add to this category of stochastic procedure: the EURO method used by EUROSTAT. The basic idea of this approach is to derive I-O tables that are consistent with official macroeconomic forecasts for GDP but avoid arbitrary adjustments of input coefficients to ensure the consistency of supply and demand. More specifically, this method should only use official macroeconomic forecasts as exogenous input for the iterative procedure. Column and row vectors for intermediate consumption and final demand should be derived as endogenous variables rather than accepted as exogenous variables from unspecified sources. The EURO method, like any other method presented above, presents advantages related to simplification of numerically and conceptually complex problems with substantial cost advantages. However, these advantages come at the price of certain disadvantages. In fact, primary forecasts for output levels normally not being available, it should be noted that the structural composition of final demand estimates in the Euro procedure is not based on econometric functions. Moreover, the impact of relative prices and other important economic variables such as innovation, technical progress, and productivity gains become difficult to fully anticipate.

3.3.2 RAS Formalism and its Limits

While the RAS method was implemented in the 1930s, Stone adapted the technique in 1961 for use in updating I-O tables from the work of Deming and Stephan (1940). The method is used when new information on the matrix row and column sums becomes available and we need to update a fully existing matrix.

Thus, following Lemelin, Fofana & Cockburn (2005) and Robinson, Cattaneo & El-Said (2001), the basic problem is generating a new *nxn* matrix of A^1 from an existing matrix A^0 of the same dimension under restriction of the new given row and column totals X and Y. We then need to apply row and column multipliers *r* and *s*, respectively. The (2n – 1) unknown multipliers are determined by the (2n – 1) independent row and column restrictions using an iterative adjustment procedure.

Suppose we need to update a social accounting matrix (SAM). If we define *T* as a SAM transactions matrix, where t_{ij} is a cell value that satisfies the next restrictions:

$$T_{\bullet j} = \sum_i t_{ij}$$

to construct the coefficient matrix $[a_{ij}] = A$ of a SAM, we divide each cell t_{ij} by the row total $t_{\cdot j}$ of a corresponding column, i.e.:

$$a_{ij} = {t_{ij}}/{t_{\bullet j}}$$

In this case, if we indicate the unknown coefficients of A^1 by $[a^1_{ij}]$ and known coefficient elements of A^0 by $[a^0_{ij}]$, the RAS procedure is as follows:

$$a^1_{ij} = r_i\, a^0_{ij} s_j,$$

In matrix notation, we have:

$$A^1 = \widetilde{R} \overset{0}{A} \widetilde{S},$$

Where $\sim$ indicates the diagonal matrix of elements R and S. Then, this last equation shows that the RAS method successfully constitutes an iterative algorithm of biproportional adjustments.

Let us indicate the serial number of different steps of the RAS algorithm by 1,2,…t and then define each algorithm as the following:

Step 1:

$$a^1{}_i = \hat{x}_{i\bullet} \Big/ \sum_j x_{ij}{}^0 \Rightarrow x_{ij}{}^1 = a^1{}_i x_{ij}{}^0 \Rightarrow b^1{}_j = \hat{x}_{\bullet j} \Big/ \sum_i x_{ij}{}^1 \Rightarrow x_{ij}{}^2 = b_i{}^1 x_{ij}{}^1,$$

Step 2:

$$a^2{}_i = \hat{x}_{i\bullet} \Big/ \sum_j x_{ij}{}^2 \Rightarrow x_{ij}{}^3 = a^2{}_i x_{ij}{}^2 \Rightarrow b^2{}_j = \hat{x}_{\bullet j} \Big/ \sum_i x_{ij}{}^3 \Rightarrow x_{ij}{}^4 = b_i{}^2 x_{ij}{}^3,$$

Step t:

$$a^t{}_i = \hat{x}_{i\bullet} \Big/ \sum_j x_{ij}{}^{2t-2} \Rightarrow x_{ij}{}^{2t-1} = a^t{}_i x_{ij}{}^{2t-2} \Rightarrow b^t{}_j = \hat{x}_{\bullet j} \Big/ \sum_i x_{ij}{}^{2t-2} \Rightarrow x_{ij}{}^{2t-2} = b_i{}^t x_{ij}{}^{2t-1}.$$

Thus, the step t corresponds to the last iteration ensuring a final convergence solution.

Compared to the cross-entropy approach presented above, the RAS procedure presents an obvious advantage of being relatively simple to use. However, it presents the following severe drawbacks:

Lack of underlying economic theory and limited possibility of constraining information

For instance, there is a need to fix some cell values inside an I-O matrix during its updating procedure when these values are known with sufficient certainty. The RAS procedure is not well suited for tackling such problems.

No possibility to treat the problem on the stochastic side

For instance, if the new known row and column totals, X and Y, are known with uncertainty—a realistic hypothesis—one will need to add a random component to the model. The RAS procedure is not suited for tackling such problems.

The risk of starting with a basic I-O matrix A^0 characterized by "spurious consistency"

This is the case when the matrix has been updated or balanced on the basis of a theoretical hypothesis, e.g., macroeconomic closure rules. In such a case, the matrix appears well balanced despite possibly containing systematic errors.

Because of the above empirical problems, many researchers have tried the extension of the RAS procedure in the hope of rendering it more flexible. Lemelin et al. (2005) show that the RAS procedure can be apprehended as a Bayesian information processing rule with the new known row and column totals X and Y taken for new data in the Bayesian sense. In the process, after having compared the Lagrange multipliers of both procedures, the authors show conditions of equivalence between the RAS procedure and the Kullback-Leibler cross-entropy approach. Similar work has been presented by McDougall (1999). He concluded that the RAS approach corresponds to a generalized Shannon cross-entropy technique, suggesting that the latter cannot supplant the former. Nevertheless, according to McDougal, the cross-entropy approach can extend and adapt the RAS technique to problems that do not fit well with the traditional matrix balancing framework. Interestingly enough, Robinson et al. (2001), have conducted a comprehensive experiment on a 1994 SAM of Mozambique. Starting from the balanced SAM and randomly imposing new row and column totals, the authors have operated a Monte Carlo experiment in which they have simultaneously updated the matrix using RAS and Shannon cross-entropy procedures. They found RAS and CE to be equivalent measures—meaning that RAS is an entropy theoretic—if the CE method uses a single cross-entropy measure as an objective instead of attempting to use the sum of column cross-entropies. They concluded by confirming the findings of many previous researchers according to whom the RAS procedure remains less flexible in the case of new information in comparison with the cross-entropy technique, which is better at processing new information for optimal consistency of the updated SAM.

However, due to its popularity, researchers have proposed new extensions of the RAS approach [e.g., http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-RA-07–013-EN.pdf].

One of the interesting extensions has been the Model of Double Proportional Patterns (MODOP) developed by Stäglin (1972:69–81). This model consists of estimating all the existing cells of transaction. The resulting inconsistent matrix is then estimated by the RAS approach. The basic idea is to calculate the geometric mean of row and column multipliers and then to apply this factor to each element of the matrix. Following the above author, outputs from the MODOP are often similar to those from the RAS approach. Some trials that allow the RAS approach to constrain targeted cells inside a matrix or to render the cells stochastic have been undertaken in recent years.

In a recent, thorough study on the comparative performance of the cross-entropy and RAS techniques, Chisari et al. (2012) concluded that cross-entropy had a more general character for the following reasons:

a) It does not need all the new totals of rows or columns (although prediction will be less accurate).
b) It does not need a balanced initial matrix (the sum of rows could be more/less than the sum of columns).
c) New rims could contain an error term.
d) New rims can be non-fixed parameters.
e) Many values on the final matrix could be fixed (not necessarily a parameter).
f) It allows non-linear constraints.

Referring to their simulation outputs, the authors propose *a rule of thumb consisting in preferring the RAS method if and only if any constraint or one constraint is enforced*. This seems to explain why the RAS approach continues to be successfully applied in different prediction studies. Furthermore, comparing the starting point to the RAS method, the above authors observe that the purchasing method is preferred to the supplying matrix because the aggregate bias is in this lower case. Furthermore, note that this suggestion does not seem to be consistent with the above investigations done by Robinson et al. (2001) on the Mozambique economy according to which the RAS and Shannon entropy approaches produce the same performance when no additional restriction is imposed. We will come back to this point when we present an example of I-O updating at the end of this section.

Trying to assess the cross-entropy approach in a dynamic, stochastic multi-objective optimisation problem, Bekker & Aldrich (2011) have concluded that acceptable results can be obtained while doing relatively few evaluations. Such an empirical fact tends to confirm a large area of cross-entropy application.

Finally, the central point to focus on through this section has been the limit—at least according to existing literature—of the RAS method compared to Shannon-Kullback-Leibler cross-entropy. Thus, since the latter is itself a converging case of Tsallis non-extensive entropy, the RAS procedure can be seen as relatively less attractive with respect to both entropy approaches.

3.3.3 Application: Updating an Aggregated EU I-O Matrix

3.3.3.1 The RAS Approach

Tables 8 and 9 represent 27 aggregated EU symmetric I-O tables for domestic output at basic prices for years 2006 and 2007. In this simplified example, let us suppose that we have no information—though we do—on the elements of the intermediate consumption sub-matrices for the period 2007 in Table 2. Using both the RAS and cross-entropy approaches, we are asked to predict those elements on the basis of the

2006 I-O matrix and sectorial accounts totals of the period 2007, which are supposed to be known.

In the example below, we directly use the transaction matrix in current price value. Thus, instead of coefficient matrix A presented in the above RAS formula, we use the matrix X of transactions. Using the formalism explained in the above section for the RAS approach, we present below the algorithm to solve the problem:

Iteration 1
Calculation of row multipliers

1.098105 1.068032 1.066862 1.044765 1.052929 1.037053

							Actual	Row multipliers
50495.53	184777.3	3012.483	24612.5	3822.238	6453.238	273173.2	280691	1.027522
84702.14	2585662	382824.5	542838.5	230943.7	266151.9	4093122	4128898	1.008741
2941.775	45081.56	361192.4	47169.66	121551.4	47356.15	625293	598576	0.957274
38756.27	770976.1	124401.9	865048.6	264081.2	185633.2	2248897	2228709	0.991023
27106.73	694435.6	18950.7	753990.3	1314546	370177.4	3349764	3236274	0.96612
5269.119	83635.91	11198.4	72001.38	108426.3	267284.2	547815.4	538038	0.982152

where, e.g., starting multiplier 1.098105 to be later multiplied by the first column elements of the initial transaction matrix of 2006 is obtained from the quotient 419340/381876, i.e., the first column output of 2007 divided by the first column output of 2006, both respectively from Tables 9 and 8.

With elements of row multipliers on the diagonal matrix premultiplied by X^0ij obtained in the above transformation, we get:

$\tilde{R}\,X^{0}_{ij}$ equal to:

1.027522	0	0	0	0	0
0	1.008741	0	0	0	0
0	0	0.957274	0	0	0
0	0	0	0.991022853	0	0
0	0	0	0	0.96612	0
0	0	0	0	0	0.982152

X

50495.53	184777.3	3012.483	24612.5	3822.238	6453.238
84702.14	2585662	382824.5	542838.5	230943.7	266151.9
2941.775	45081.56	361192.4	47169.66	121551.4	47356.15
38756.27	770976.1	124401.9	865048.6	264081.2	185633.2
27106.73	694435.6	189507.7	753990.3	1314546	370177.4
5269.119	83635.91	11198.4	72001.38	108426.3	267284.2

equal to:

	51885.25	189862.6	3095.392	25289.88	3927.432	6630.843
	85442.48	2608262	386170.6	547583.2	232962.2	268478.2
	2816.084	43155.38	345759.9	45154.27	116358	45332.79
	38408.35	764055	123285.2	857282.9	261710.5	183966.7
	26188.35	670908	183087.1	728445	1270009	357635.7
	5175.075	82143.16	10998.53	70716.28	106491.1	262513.7
	209915.6	4358386	1052397	2274472	1991459	1124558
Actual	210592	4373650	1056144	2280581	1974301	1115919
Column multiplier	1.00322	1.003502	1.003561	1.002686	0.991384	0.992318

$X^1 = \widetilde{R}\overset{0}{X}_{ij}\widetilde{S}$ equal to:

51885.25	189862.6	3095.392	25289.88	3927.432	6630.843
85442.48	2608262	386170.6	547583.2	232962.2	268478.2
2816.084	43155.38	345759.9	45154.27	116358	45332.79
38408.35	764055	123285.2	857282.9	261710.5	183966.7
26188.35	670908	183087.1	728445	1270009	357635.7
5175.075	82143.16	10998.53	70716.28	106491.1	262513.7

X

1.00322	0	0	0	0	0
0	1.003502	0	0	0	0
0	0	1.003561	0	0	0
0	0	0	1.002685984	0	0
0	0	0	0	0.991384	0
0	0	0	0	0	0.992318

equal to:

52052.34	190527.6	3106.414	25357.81	3893.595	6579.903
85717.63	2617396	387545.8	549054	230955.1	266415.7
2825.153	43306.52	346991.2	45275.55	115355.5	44984.53
38532.04	766730.8	123724.2	859585.6	259455.7	182553.5
26272.69	673257.6	183739.1	730401.6	1259068	354888.3
5191.74	82430.84	11037.69	70906.23	105573.6	260497

Iteration 2. ...j....:
Repeat Iteration 1 algorithm
In this example, we reach the optimal value at the seventh iteration and then get the following values of the new supposed unknown transaction matrix of the year 2007:

0.999998	0	0	0	0	0
0	0.999999	0	0	0	0
0	0	1.000001	0	0	0
0	0	0	0.999999944	0	0
0	0	0	0	1.000001	0
0	0	0	0	0	1.000006

X

51929.5	190003.4	3095.379	25249.69	3870.157	6543.312
85635.77	2613871	386712.9	547482.9	229888.1	265307.2
2829.519	43356.44	347112.1	45258.99	115109.9	44909.49
38562.37	767033.6	123673.6	858620.7	258707.4	182111
26405.53	676396.6	184447.4	732693.9	1260792	355538.3
5228.733	82985.65	11103.05	71275.19	105935.8	261511.4

X

1.000001	0	0	0	0	0
0	1.000001	0	0	0	0
0	0	1	0	0	0
0	0	0	0.999999724	0	0
0	0	0	0	0.999999	0
0	0	0	0	0	0.999998

equal to:

51929.54	190003.5	3095.379	25249.68	3870.153	6543.302
85635.84	2613873	386712.9	547482.7	229887.9	265306.8
2829.521	43356.47	347112	45258.98	115109.8	44909.42
38562.4	767034	123673.6	858620.4	258707.2	182110.7
26405.55	676397	184447.4	732693.7	1260790	355537.7
5228.738	82985.69	11103.05	71275.17	105935.7	261511

Table 6: Error(%) prediction from the RAS procedure

3	0	1	-4	-7	4
1	0	0	-2	0	0
-7	5	-1	6	-4	3
-6	-1	-2	1	0	3
1	0	2	0	0	-1
-4	1	4	2	2	-2

The above errors are calculated as the error discrepancy percentage between the true matrix of transaction representing the period of year 2007 and the matrix updated by the RAS procedure on the basis of the 2006 I-O matrix.

3.3.3.2 The Entropy Approach

For comparative purposes, let us apply entropy formalism for updating the same IO transactions table as in the above example related to the RAS approach. Thus, using the Tsallis entropy formalism presented in (3.17) and in (2.48–2.50) under the hypothesis that transaction totals of the targeted period are known (and without any additional restriction), we get the outputs presented in Table 7.

Comparison of Tables 6 and 7 shows slightly higher performance of the RAS approach. This seems to confirm the rule of thumb proposed by Chisari et al. (2012): *if and only if any constraint or one constraint is enforced* as in the present case. However, such a conclusion, as already mentioned above, is not in line with the one proposed by investigations conducted by Robinson et al. (2001) which lead to equivalent performance in the same conditions between the cross-entropy and RAS approaches. More investigations are needed to contradict or confirm the study results of the above authors. Of course, following the results of several investigations presented above, cross-entropy naturally presents higher performance than the RAS approach, particularly when statistical data are known with uncertainty. In the next section, we are going to propose the forecasting of a higher dimension I-O table through cross-entropy formalism. Later, when treating the case of a SAM with higher dimension, we

Table 7: Error (%) prediction of the I-O table by the Tsallis (or Shannon) cross-entropy procedure

Products (CPA)	P1	P2	P3	P4	P5	P6
P1	7.985	3.141	1.934	-0.049	-7.602	4.579
P2	8.954	2.214	-0.722	-0.099	-2.772	-0.158
P3	-8.035	2.042	-5.767	3.010	-11.670	-2.940
P4	-4.044	0.691	-4.533	7.906	-4.444	0.523
P5	0.370	-2.219	-2.198	-1.980	-6.118	-0.147
P6	-2.688	0.587	1.827	1.848	-2.704	-4.190

Table 8: Symmetric I-O Tables for domestic output at basic prices (year 2006; EU27, Mio. EUR current prices)

	Inputs of products							
Products (CPA)	P1	P2	P3	P4	P5	P6	Others	Total
P1	45984	173007	2824	23558	3630	6223	126650	381876
P2	77135	2420959	358832	519579	219335	256642	2999462	6851944
P3	2679	42210	338556	45149	115441	45664	1130576	1720275
P4	35294	721866	116605	827984	250806	179001	2465270	4596826
P5	24685	650201	177631	721684	1248467	356951	1932724	5112343
P6	4798	78308	10497	68916	102976	257734	3054393	3577624
Others	191301	2765392	715330	2389956	3171688	2475408	1499777	13208852
Total	381876	6851944	1720275	4596826	5112343	3577624	13208852	

Source: own calculations.
Source: based on http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/overview

will present the methodology of updating it in the presence of uncertainty and with a free number of restrictions.

3.3.3.3 I-O table Forecasting

By "updating," we compute operations on table rows and columns with the purpose of balancing its row and column totals, but a forecasting operation implies the use of a certain theoretical model—whether deterministic or not. To forecast an I-O table, the statistical data concerning final demand Y_i and the value added D_j are generally available or could be obtained on the basis of existing information. For example, in the case of Eurostat, this information exists for the time horizon year 2013 while

Table 9: Symmetric I-O Tables for domestic output at basic prices(year 2007)

	Inputs of products							
Products (CPA)	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**Others**	**Total**
P1	53529	189315	3126	24311	3626	6784	138648	419340
P2	86528	2624065	386717	535924	229362	266302	3189197	7318095
P3	2656	45671	344941	48062	111100	46147	1236719	1835296
P4	36335	756877	121086	869135	258074	187202	2573895	4802604
P5	26539	673893	188669	730656	1264384	352132	2146659	5382932
P6	5005	83827	11606	72494	107755	257351	3172149	3710187
Others	208748	2944446	779152	2522023	3408631	2594266	1599374,605	14056642
Total	419340	7318096	1835296	4802604	5382933	3710185	14056642	

Source: based on http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/overview

where :

P1	Products of agriculture, hunting and fishing
P2	Industrial products (incl Energy)
P3	Construction work
P4	Trade, transport and communication services
P5	Financial services and business services
P6	Other services
Others	Use of imported products, cif
	Taxes less subsidies on products
	Value added at basic prices
Total	Row or column totals

neither the value of global product X_j nor the matrix of cost structure $[a_{ij}]$ are known for that forecasting period. Here, the last updating of IO tables concerns the period 2007. Building an I-O table is time consuming and a lot of information, particularly concerning the transaction matrix (including import elements), is not easy to assess. Economic information from different enterprises or industries must be updated on the basis of new flows of additional information. As a result, getting a final version of that table comes long after with many lag years. Under such conditions, finding a workable methodology for setting up a robust prediction technique of an I-O table should bear precious advantages.

Therefore, the key question is: Based on the I-O table of the previous period, on vectors $Y^p{}_i$ and $D^p{}_j$ (both from the forecasting period), is it possible, using the connection (3.2), to estimate the unknown I-O table^p of the forecasting period? Of course, the answer is not affirmative if we use the classical statistical-mathematical approach to solve that inverse problem. On the one hand, we do not have information about matrix A^p of the forecasting period to derive the value X^p using relation (3.2) and, in this way, to determine I-O table^p. On the other hand, as an effect of the possibility of estimating X^p, for example using a dynamic investment model, the new question that arises is: Disposing of X^p and of Y^p of the forecasting period, could one possibly determine the matrix coefficients A^p on the basis of the same relation (3.2)? Since the IO table has size NxN, and we have by assumption only information about final demand Y^p and global product X^p of each sector, this means that we have $(N - 2)x(N - 2)$ degrees of freedom, where N is, once again, the number of branches in the IO table. Such a problematic belongs then to the category of inverse problems, which suggests that there may be an infinity of matrices A^p that could reproduce the identical values of final demand Y^p and the global product X^p. Among them, we will choose the one that best maximizes consistency information between the prior, data, and the posterior. We can retrieve a solution proposal to that problem in the second part of this monograph about the maximum entropy principle or relative (cross) entropy. Let us again present it below in the context of updating and forecasting an I-O table on any other form of its extension.

3.3.3.4 The Non-Extensive Cross-Entropy Approach and I-O Table Updating

As suggested above, in the recent literature there are several methods for updating and balancing elements of national accounts balance sheets, for instance, an I-O table when equality of corresponding sums of columns and rows is required. Some of their limits have been emphasised here. Preference is then given to methods based on statistical theory of information for their capacity to adjust information when structural changes affect the economy or when additional consistent information has to be added to what already exists. The most frequently used theoretic-information methods are the Maximum Likelihood, Bayesian method of moments, and methods based on the maximum entropy principle. Through their article, Giffin & Caticha (2007) have proven that the principle of maximum entropy represents a generalization of the Bayesian approach as a method of inference on the basis of an *a priori* information. Probably for these reasons, application of the cross-entropy approach to balance the social accounting matrix has been widely adopted in empirical application during recent years (e.g., Robinson et al., 2001).

As we know from Part II of the monograph, on the basis of the results of Shannon (1948) and Jaynes (1957), Kulback-Leibler (KL) (1951, 1957) and Good (1963) have proposed the principle of minimum (relative) entropy. This principle aims at assessing *a posteriori* parameters (probabilities P) of the most plausible, shortest divergence in relation to *a priori* parameters (probabilities Q), under restrictions related to data moments, normalization condition, or any other *a priori* information presenting consistency with probabilities in the criterion function. The formulation in the case of discrete events is like in (3.16) and, thus, we have:

$$\text{Min}(r_j, r_0) \to I_q\left[r \| r_o\right] = \frac{1}{q-1}\sum_j r_j\left[(r_j)^{q-1} - (r_{0j})^{q-1}\right] - \sum_j (r_j - r_{0j})(r_{0j})^{q-1} \qquad (3.16')$$

or, as traditionally done in the case of Shannon-Gibbs entropy,

$$Min_{KL}(P\|Q) = \sum_{i=1}^{n} p_i \ln\left({p_i}/{q_i} \right) = P' \ln P - P' \ln Q, \text{ for } i = 1,\ldots,n$$

under restrictions:

$$X_j = \sum_{j=1..n} \frac{(r_{.j})^q}{\sum_{j=1..n}(r_{.j})^q} X_j \qquad (3.18)$$

$$\sum_{j=1..n} D_j \frac{(r'_j{}^d)^q}{\sum_{j=1..n}(r'_j{}^d)^q} = \sum_{i=1..n} \frac{\left(r_i^{y'}\right)^q}{\sum_{i=1..n}\left(r_i^{y'}\right)^q} Y \qquad (3.19)$$

$$\Omega(h) = F \qquad (3.20)$$

$$P'I = 1 \text{ and } P \geq 0. \qquad (3.21)$$

We adopt Shannon-Gibbs symbols in the criterion function above. Correspondence with the Tsallis criterion function in (3.16) is easy to do. The symbol P corresponds to r and Q to ro. Matrix P stands for posterior probabilities guaranteeing the balance of a previously unbalanced table, the elements of which sum up to unity by column (sector). Q is a matrix of coefficients from a known table. In the case of the I-O table (see Table 2) elements of Q are derived by dividing each column element by its total. They then represent input coefficients except the case of the final demand column, where these coefficients explain the structure of product consumption for a given final demand institution. Thus, its elements must satisfy the additivity condition. Equation (3.18) demonstrates that the column total must match with corresponding row elements multiplied by corresponding probabilities (coefficients) P_j. Equation (3.19) states equality between value added and final demand totals, with $p'_j{}^d$ and $P_i^{y'}$ being transposed respective vector of sectorial value added components and vector of structural coefficients of final demand. Probabilities are presented in escort distribution formulation presented in footnote 19. Functional h in (3.20) gives a piece of information which has a significant relationship (consistency) with probabilities in the criterion function. This may be, for example, a macroeconomic balance equation or any distribution of a treatment of errors. Equation (3.21) is one of the additivity conditions of probabilities and reminds us that no probability can take a negative value.

3.3.3.5 Forecasting an I-O Table

Let us first formulate the model. As previously stated, a fundamental problem results from the lack of information about sectorial global product X_j^p or equivalent input transactions x_{ij} since we assume final demand to be known. Thus, before implementing entropy formalism (3.16), we must first estimate the values X_j^p. For this purpose, we propose using a dynamic model[24] of I-O in which investment is the endogenous variable in the context of accelerator analysis of macroeconomic theory. According to macroeconomic theory, it is expected that investment is induced if final demand is expected to grow. Based on the preceding period, traditional assumptions of such a model are as follows:

- the investment is a function of the expected growth,
- information about the coefficients of material and production factors are available,
- information about capital ratios are available,
- all economic sectors are in full effect,
- capital has an infinite lifetime.

We then have the following dynamic equation of global product:

$$X^{t+1} = B^{-1}[(I - A + B)X^t - Y^t] \qquad (3.22)$$

and

$$Inv^t = B(X^{t+1} - X^t) \qquad (3.23)$$

where:

X^{t+1}: global product,
Inv^t: induced investment,
B: coefficients of capital production,
A: coefficients of material,
Y^t: final demand,
t: the index of the time.

It should be added that value of global product derived in this way does not constitute the ultimate result of the whole process of forecasting. This value only provides the *a priori* information in terms of a Bayesian viewpoint, and it will be changed in the process of entropy minimization, as previously explained. At the end of the procedure, applying relation of relative entropy in (3.16), we upgrade I-O coefficients for the period T + 1 to obtain a new post entropy I-O table. Thus let us consider a known I-O matrix BT of period T in the form of Table 2.1 displaying coefficient structure, obtained

24 See EUROSTAT.

after dividing each column element by the corresponding total column. The problem is retrieving new matrix B_{T+1} of period T + 1 for which we already know global product from (3.22) and sectorial aggregates of final demand. As is often the case in real life, let us suppose sectorial value added is known, too. If our I-O table has a dimension n x n (with n >> 2), then we have (n – 1) (n – 2) degree of freedom. Building on the formulation (2.47–2.50) involving an inverse problem, we obtain:

$$\text{Min}(b_1, b_0) \rightarrow I_q\left[B_1 \| B_0\right] = \frac{1}{q-1}\sum_i\sum_j b_{1ij}\left[(b_{1ij})^{q-1} - (b_{0ij})^{q-1}\right] - \sum_i\sum_j (b_{1ij} - b_{0ij})(b_{0ij})^{q-1} \quad (3.24)$$

under restrictions:

$$X_{1i} = \sum_{i=1..n}\frac{(b_{1ij})^q}{\sum_i (b_{1ij})^q} X_{1j} \quad (3.25)$$

with i=j in the case X1,

$$\sum_{j=1..n} D_j \frac{(b'_{1ij}{}^{d})^q}{\sum_{j=1..n}(b'_{1ij}{}^{d})^q} = \sum_{i=1..n}\frac{\left(b_{1ij}{}^{y'}\right)^q}{\sum_{i=1..n}\left(b_{1ij}{}^{y'}\right)^q} Y \quad (3.26)$$

$$\sum_{i=1..n} b_{1ij} = 1 \quad (3.27)$$

$$\Omega(h) = F \quad (3.28)$$

$$b_{1ij} \geq 0. \quad (3.29)$$

In the present model $B_t[b_{ijt}]$ with (i = 0 and 1) is a matrix of I-O coefficients for the period T_t. Thus, B_0 is *a priori* known, and we seek to estimate B_1, that is, for t = 1. Since both coefficients display probability properties of continuously belonging to interval zero-one, of summing up to unity (by column), and of additivity, we do not need to reparameterize B, and it will be taken for probability. Tsallis complexity index q informing about departure from Gaussian to power law distribution appears in the criterion function (3.24) and in constraints (3.25) and (3.26) under formulation of escort distribution, as earlier explained in footnote 19. Index one on variable X_{1j} (explaining "global product") in (3.25) refers to the period of forecast. It is equal to the vector column of coefficients multiplied by the total column, taking then into account weights related to escort distribution. A total obtained this way is equal to the total corresponding line X_{1i}. This applies equally well to (3.26), which means that totals of value-added D_j and of final demand Y_i are equal. The rest of the restrictions, i.e., (3.27–3.29) are as in (3.20–3.21).

3.3.4 Application: Forecasting the Aggregated 27 EU IO Coefficients

In this example, we will present two cases where we apply the non-extensive entropy principle[25] to forecast the aggregated 27 EU I-O coefficient table for the period 2007. The basis of that forecasting is the I-O table of the previous year, 2006, and some information related to the targeted I-O coefficients of 2007. Both tables can be found at the Eurostat statistics site: http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/overview. Since we already know the true values[26] of the I-O table of 2007, it will be much easier to assess the performance of the applied technique.

In the first case, besides information from the 2006 table, we are supposed to additionally have data on sectorial value added and final demand of the targeted period 2007. In this example, sectorial global product is neither known nor assessed. We then consider here a case which is more unfavourable than it appears in empirical investigations where, in general, many values of the targeted I-O table are assessable on the basis of existing theory or through intuition.

In the second case, besides the information used in the first case, we are supposed to additionally have the possibility of estimating some reliable inter-sector transactions of the targeted period 2007 and to incorporate them into the model as new data, according to the Bayesian model. Then, the following list of accounts which have been arbitrarily selected from the true 2007 table and information related to them have to be incorporated into the entropy model:

tobac	Tobacco products
pulpap	Pulp, paper, and paper products
Insur	Insurance and pension funding
compserv	Computer and related services
uranores	Uranium plus ores

Among the 60 accounts in the table—we have a NACE A60 breakdown—these five accounts represent 5% of all accounts and around 2% of their transaction total. Thus, in spite of using that additional piece of information, we are still dealing with an inverse problem consisting of recovering (n – 6) x (n – 6) coefficients of the I-O table on the basis of the coefficient table from the previous period, 2006, and some additional, random information from the forecasted period. Not included[27] in this book,

25 In this study, outputs from Shannon and Tsallis entropies are similar.

26 As can be verified at the electronic address http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=naio_15_agg_60_r2&lang=en, this version of the 2007 table does not yet seem to be definitive since some accounts are not balanced. The implication is that input-output coefficients related to that table cannot be considered known with certainty. At the same electronic address, we can find input-output tables of other periods, such as for 2006. However, since we have made small changes to these two tables, they will be displayed in the annex.

27 Due to lack of space, these tables cannot be presented in this book. Nevertheless, they are here

two output tables allow for assessing the performance of the proposed technique. The first table, F1, displays the precision error for the ex post forecasted inter-sector transactions for the period 2007 and is related to a minimum of hypotheses, that is, the knowledge of sectorial values added and final demand. Thus, that table displays the forecasted I-O transaction table for the period 2007 with the minimum *a priori* information. The second table, F2, displays the precision error for the forecasted inter-sector transactions of the period 2007 in the second hypothesis, that is, with some knowledge on inter-sector transactions of the five above listed accounts.

What could we learn from these outputs? Here, we will limit our comments to general aspects of cross-entropy formalism. More details will be provided in the coming chapter concerning the procedure of updating a social accounting matrix, as a generalized case of an I-O table. Table F1, as already mentioned, shows the degree of precision of the entropy procedure in forecasting the I-O coefficients of the period 2007. The only *a priori* information remains the I-O transaction matrix of 2006, sectorial values added and final demands of the period 2007. All computations have been carried out with the GAMS code. To measure the precision of the I-O coefficient, we have used the next average error variance coefficient (AEVC):

$$AEVC = \frac{\sum_i \sum_j (a_{ij} - \hat{a}_{ij})^2}{(n-1)(n-1)}$$

The coefficients a_{ij} represent the true coefficients displayed by the 2007 table of the I-O matrix prepared by Eurostat. Though the table is not fully balanced, we still consider it sufficient to correspond to a true data generating system. The forecasted coefficients $\hat{a}_{ij}$ are supposed to be affected by a certain margin of error. The denominator explains the degree of freedom of a Warlasian equilibrium table, i.e., the number of accounts[28] but one.

Let us first consider a simple benchmark measure of the information divergence between the 2006 and 2007 tables. In these circumstances, values of coefficients $\hat{a}_{ij}$ represent the period 2006 and values of the coefficients a_{ij}, that of 2007. Thus, the derived variance value AEVC is 0.00003846. In the context of information theory, this value corresponds to the minimum entropy result when any restriction (except normality conditions) to the criterion function has been enforced. It corresponds, too, to the maximum entropy outputs in the same conditions.

Now, if we consider the first case when we know the inter-sector values added and final demands, we get a new AEVC equal to 0.00002932 and representing a discrepancy between the true and the estimated values. In comparison with the benchmark coefficient, the new piece of information has brought about an overall coeffi-

numbered F1 and F2 so that the interested reader can, on request, get a copy of them from the author.

28 For computational reasons, two accounts concerning the mining sectors have been aggregated into one account.

cient variance reduction of around 24%. Though this coefficient is very small, a look at the non reported in this book outputs (table “F1") could reveal a high variation in these coefficients. As it has been noticed in Shannon entropy econometrics literature, entropy formalism tends to shrink small probabilities in favour of those higher. Through this example, we discover the same stylised fact found in the case of Tsallis entropy formalism. Inside that table, we notice the presence of a 100% shift of many values. This is the result of small probability shrinkage, as explained above. To reduce these variations, we need to impose additional restrictions on the model. On theoretical grounds, the formal causes of small probability shrinkage are well presented in the work (see previous chapters) of Golan et al. (1996) as an extension of the work of Greenberg et al. (1989) on the family of Stein-rule estimators[29] proposed by Stein & James (1961).

In the last experiment carried out, we additionally consider that some transaction values of the above listed accounts are known. We then get a new AEVC equal to 0.00001042502. In comparison with the benchmark coefficient, the new piece of information has brought about an overall coefficient variance reduction of around 74%. Noting that the small variance AEVC naturally represents the average for all values inside the table, this seems to contrast with the relatively higher precision error affecting many cell values of the tables F1 and F2. In general, as underscored above, higher imprecisions tend to affect smaller values (then with smaller weights) inside the table cells.

3.3.5 Emission Coefficients Forecasting: A Theoretical Model

One may now combine the two previous procedures of retrieving information in the case of the inverse problem. The first procedure has assessed emission coefficients when sectorial global product and the transactions matrix were available (see section III.3.2.2). The second has assessed the I-O table on the basis of knowledge of sectorial final demand of current period and *a priori* information about the table of preceding period. Now, the interesting problematic could be assessing sectorial emission coefficients on a knowledge basis of:

- total emission by theme of the forecast period,
- sectorial final demand of the forecast period,
- an input-output table of a recent period.

29 The main idea is that if three or more unrelated parameters are measured using the James-Stein estimator, their total square error will be lower than in the case of the least square (LS) estimator well known to provide the lowest variance among all other linear estimators. However, when each parameter is estimated separately, the LS estimator leads to higher performance. This is so because of the tendency of the Stein family of estimators to shrink small probabilities of the estimated system.

Table 10: A Simplified Environmentally extended I-O table

i	Sector	Flows					Final demand		Total
1	X_1	x_{11}	x_{12}	x_{13}	…	x_{1m}	Y_1	0	
2	X_2	x_{21}	x_{22}	x_{23}	…	x_{2m}	Y_2	0	X_2
…	…	…	…	…	…	…	…	0	…
m	X_m	x_{m1}	x_{m2}	x_{m2}		x_{mm}	Y_m	0	X_m
m +1	D_j Value-added	D_1	D_2	D_3	…	D_n	0	0	D
…	E_j	e_{11}	e_{12}			e_{1m}	0	e_{Y1}	E1
…	…	…	…	…	…	…	0	…	…
n	(Energy themes)	e_{f1}	…	…	…	e_{fm}	0	e_{Yf}	E_f
	Total	X_1	X_1	X_1	…	X_m	Yv	Ye	Total

Source: own elaboration.

Solution of such an inverse problem can be set up as follows:

$$\text{Min} \rightarrow I_q\left[p \| p_0\right] = \frac{1}{q-1} \sum_j p_j \left[(p_j)^{q-1} - (p_{0j})^{q-1}\right] - \sum_j (p_j - p_{0j})(p_{0j})^{q-1} \tag{3.17'}$$

subject to

$$X_{1i} = \sum_{i=1..n} \frac{(p_{1ij})^q}{\sum_i (p_{1ij})^q} Z_h X_{1j} \tag{3.32}$$

with i=j in the case X_1,

$$\sum_{j=1..n} D_j \frac{(p'^{\,d}_{1ij})^q Z_h}{\sum_{j=1..n} (p'^{\,d}_{1ij})^q} = \sum_{i=1..n} \frac{\left(p_{1ij}^{\;y'}\right)^q Z_h}{\sum_{i=1..n} \left(p_{1ij}^{\;y'}\right)^q} Y \tag{3.33}$$

$$\sum_{j=1}^{n} \sum_{h=1}^{M} p_{1jh} = 1$$

$$\text{E}_\text{i} = \sum_{j=1}^{n} \sum_{h=1}^{M} \frac{p_{1jh}^{\;q}}{\sum_{j=1..n} p_{1jh}^{\;q}} Z_h X_{1j} \tag{3.34}$$

$$\sum_{m=1}^{h} e_{fm} = 1$$

with h=f + 2, i.e., additional columns due to the demand of emissions by institutions (one of them having zero value since demand of emissions must have a separate column).

All symbols are as before. Note that in this case the environmentally extended I-O matrix has the form presented in Table 10 above. Input coefficients are derived from global column totals, including, thus, quantities (instead of values) representing emissions.

3.4 Conclusions

As suggested above, generally, new information should render more homogeneous divergence between the true coefficients and those forecasted. According to the above results, we observe that new restrictions added to the model lead to a significant reduction of errors. This is in accordance with the rule of thumb proposed by Chisari et al. (2012) in section III.3.3.1 about particular conditions explaining the superiority of entropy approaches over the RAS technique.

The obtained AEVC coefficients naturally present a random character and different experiments would produce different values. However, as expected through Bayesian formalism, new data evidence will always tend to reduce the level of uncertainty or entropy.

The true difficulty in assessing a new methodology to assess a complex information system—like the one represented by an I-O table—is that, due to instruments of measure or adopted economic hypotheses, a part of the observed data may not be accurate. This can be even worse in the case of general equilibrium systems in which the balance of the whole system—or accounts—may be more or less forced. This observation is particularly true in developing countries where statistical information management can be more challenging. What we intend to explain here is that, faced with such circumstances, the output performance of the non-extensive entropy approach should, consequently, be taken with a certain margin of error. The quality of priors and model data will always play a central role.

Bibliography – Part III

Almon, C. (2000). Product-to-product tables via product-technology with no negative flows. *Economic Systems Research, Vol. 12, No.1,, 12*(1), pp. 27–43.

Avonds, L. (2007). *The input-output framework and modelling assumptions: considered from the point of view of the economic circuit.* The 16th International Input-Output Conference. Istanbul: International Input-Output Association.

Bacharach, M. (1971). *Biproportional matrices and input-output change.* Cambridge: Cambridge University Press.

Bekker, J., & Aldrich, C. (2011). The cross-entropy method in multi-objective optimisation: An assessment. *European Journal of Operational Research,* 211(1):112–121.

Beutel, J., & De March, M. (1998). Input-output framework of the European System of Accounts (ESA 1995). Paper presented at the *Twelfth International Conference on Input-Output Techniques.* New York: International Input-Output Association.

Caticha, A., & Giffin, A. (2007). Updating probabilities with data and moments. *Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering.* NY, USA: Department of Physics, University at Albany–SUNY.

Chisari, O.O., Mastronardi, L.J., & Romero, C.A. (2012). Building an Input-Output Model for Buenos Aires City. *MPRA* (40028).

Deming, W.E., & Stephan, F.F. (1940). On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics,* 11(4):427–444.

Eurostat. (2002). *The ESA 1995 Input-Output Manual - compilation and analysis (draft).* Luxembourg.

Giffin, A. & Caticha, A. (2007). *Updating Probabilities with Data and Moments*. Presented at the 27th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Saratoga Springs, NY, July 8–13, 2007.

Gosh, A. (1964). *Experiments with Input-Output Models, An Application to the Economy of the United Kingdom, 1948–1955.* Cambridge: Cambridge University Press.

Greenberg, E., Denzau, A.T., & Gibbons, P.C. (1989). Bayesian Estimation of Proportions with a Cross-Entropy Prior. *Communications in Statistics - Theory and Methods,* 18(5):1843–1861.

Lemelin, A., Fofana, I., & Cockburn, J. (2005). *Balancing a social accounting matrix: Theory and application.*Centre Interuniversitaire sur le Risque des Politiques Economiques et l'Emploi (CIPREE), Université de Laval.

Leontief, W. (1941). The Structure of American Economy, 1919–1929. New York: Cambridge, (Mors): Harvard University Press, (Second Ed. 1951, Oxford University Press).

Leontief, W. (1966). *Input-Output Economics.* New York: Oxford University Press.

Leontief, W. (1970). The Dynamic Inverse. *Contributions to Input-Output Analysis*, pp. 17–46.

Leontief, W. (1986). *Input-Output Economics, Second Edition.* New York: Oxford University Press.

McDougall, R.A. (1999). Entropy Theory and RAS are Friends. *Working Paper No. 6, GTAP*.

Miller, R., & Blair, P. (1985). *Input-output analysis: Foundations and extensions.* Englewood Cliffs, N.J.: Prentice-Hall.

Robinson, S., Cattaneo, A., & El-Said, M. (2001). Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods. *Economic Systems Research,* 13(1):47–64.

Snower, D.J. (1990). New methods of updating input-output matrices. *Economic System Research,* 2:27–38.

Stäglin, Reiner (1972), MODOP - Ein Verfahren zur Erstellung empirischer Transaktionsmatrizen, in: Münzner, H. and Wetzel, W.: *Anwendung statistischer und mathematischer Methoden auf sozialwissenschaftliche Probleme*, Würzburg, pp. 69–81.

Stein, C., & James, W. (1961). Estimation with quadratic loss. *Proc. Fourth Berkeley Symp.*:361–379).

Stone, R. (1955). Input-Output and the Social Accounts. *Proceedings of an International Conference on Input-Outputs Analysis.* New York: University of Pisa.

Stone, R. (1981). Input-Output-Analysis and Economic Planning: A Survey. *Mathematical Programming and its Applications.* Milan: Angeli.

Stone, R. (1984). Balancing the National Accounts: The Adjustment of Initial Estimates. *Demand, Equilibrium and Trade*.

Toh, M.H. (1998). The RAS approach in updating input-output matrices: An instrumental variable interpretation and analysis of structural change. *Economic Systems Research,* 10(1):63–78.

Tomaszewicz, Ł. (1992). *Przepływy miedzygałeziowe. Elementy teorii, Wydawnictwo UŁ (rozdz. 1–3, 7,)*. Łodż: University of Łodż.
Tomaszewicz, Ł. (2005). Metody analizy input-output. *Przegląd Statystyczny, 52*(4), strony 15–22.
United Nations. (1993). *Integrated Environmental and Economic Accounting, Handbook of National Accounting*. New York, USA: United Nations.
Eurostat: http://ec.europa.eu/eurostat/web/esa-supply-use-input-tables/overview.

PART IV: **Social Accounting Matrix**

1 Position of the Problem

After presentation, in the previous part of an input-output table in the context of ill-behaved inverse problem solution, let us talk about the next national accounts table, known as a Social Accounting Matrix (SAM) (e.g., Stone, 1970; Graham, 1985; Scandizzoa & Ferrareseb, 2015). From a macroeconomic point of view, this table is a generalization of an input-output table. While input-output shows primary distribution of income without saying anything about its secondary distribution, the SAM table fills in this missing information and, at the same time, displays a complete flow of products and values in a general equilibrium framework. In the coming section, the ill-behaved inverse problem will be particularly addressed since SAM table building, more than for input-output tables, requires more information to be gathered from different sources but involving, consequently, higher risk of statistical inconsistencies when this information is aggregated. In this chapter, we present the general aspects of a SAM by introducing the parameter or multiplier estimation approach, and we then discuss the limits of classical econometric methods.

For the last two decades, Kullback-Leibler minimum entropy (KLME) formalism has encountered relative success in the social sciences, particularly when a solution was required for inverse problems (e.g., Robinson et al., 2001; Bwanakare, 2013). The objective of the present document is to extend the KLME approach to the non-ergodic Tsallis entropy system, represented in the present study by an initially non-balanced quadratic social accounting matrix (SAM), known to display Walrasian general equilibrium features.

The described economic system is then defined by different interactive subsystems, each represented by respective actors and characterized by optimizing behaviour. Households, which tend to maximize a certain utility function, remain the owner of factors of production and are the final consumer of produced commodities while firms maximize profits by optimal renting of these factors from households for the production of goods and services. In this model, government has the passive role of collecting taxes and disbursing tax revenue. Furthermore, the economy analysed is small and open and a ‘price taker’ from the rest of the world. The above optimal behaviour inside subsystems leads to general market equilibrium in all respective sectorial markets.

A SAM table is a statistical tool to summarize all the above economic transactions by registering, in their respective rows and columns income and expenses in accordance with the double-entry book-keeping principle. However, due to different and sometimes contradictory sources of collected statistical information, the SAM is not balanced, i.e., respective column or row totals are not matching. Such statistical data may display, as partially coming from statistical surveys, systematic errors, most of the time evidenced through a tail queue Gaussian distribution.

Since a SAM-based model contains more unknown parameters to estimate than the number of determined equations, updating and balancing such stochastically unbalanced matrices belongs to the category of the generalized inverse problem.

2 A SAM as a Walrasian Equilibrium Framework

In this part of the present work, we are going to present different macroeconomic aspects of a SAM in the context of further macroeconomic models to be presented and estimated. It is useful to remind readers about Walrasian general equilibrium features of a SAM. This will increase understanding of its construction and improve interpretation of its post-estimation outputs. This point of interpreting outputs will appear with more acuity in the coming sections when, at the end of the estimation process by the maximum entropy principle, we have to interpret parameters of the estimated model. Though in the next chapter on the computable general equilibrium model we will examine the philosophical underpinnings of the SAM in terms of economic theory, we must immediately be aware that both the input-output table and the SAM were conceived as practical applications of the general equilibrium theory earlier introduced by Walras, one of initiators of the Australian school of thought. Brown & Stone (1962) has provided a definition of Walrasian hypotheses as follows:

> H1. *Observed market demand is the sum of consumers' demands derived from utility maximization subject to budget constraints at observed market prices.*

> H2. *There exists an observable (locally) unique equilibrium price system such that the observable market demand is equal to the observable market supply in every market.*

> H3. *The observed equilibrium price system is a (locally) stable equilibrium of trial-and-error price adjustment.*

The first hypothesis fixes the prerequisites under which Walrasian equilibrium is feasible. The second and third hypotheses specify quantitative relations which lead to equilibrium. Equilibrium in the economic flows results in the conservation of both product and value (Liossatos, 2004).

Additionally, the three conditions of *market clearance*, *zero profit*, and *income balance* are employed by CGE modellers to solve simultaneously for the set of prices and the allocation of goods and factors that support general equilibrium. In terms of circular flows in economy, each row total of each economic sector is equal to a corresponding column total, and in that way a general equilibrium of macroeconomic aggregates is ensured.

Research contributions in this area are, to our knowledge, very limited. In their pioneering work, Duncan & Smith (Duncan, 1999; Liossatos, 2004) found that Walrasian equilibrium is not guaranteed by a free market existence. Without implausibly strong restrictions on the production sets and preferences (for example, that the production sets do not exhibit increasing returns to scale, a pervasive feature of real technologies), the demand and supply correspondences may be empty for some prices and may be discontinuous, so that no equilibrium price system can be found.

The question of the existence and stability of a Walrasian equilibrium encounters mathematical difficulties and paradoxes. In particular, the question of finding a robust stability in equilibrium prices has remained elusive, and the issue of the existence of Walrasian equilibrium has been settled only by introducing into the argument powerful abstract mathematical principles, which have no real economic foundation (see, e.g., Duncan & Smith, 2009). Taking the above into account, one may think that Walrasian equilibrium, at least on empirical grounds, is a kind of approximation of the Pareto efficiency benchmark.

Using the maximum entropy principle, the above authors (Duncan, 1999; Liossatos, 2004) have recently tried to prove the existence, the uniqueness, and the stability of the Pareto optimum.

The starting hypothesis is to consider

> a set of feasible market transactions as typically large, that is, once the number of types, the number of traders of each type, and the number of points in the offer sets become moderately large. Furthermore, there are many different ways of assigning traders to transactions in their offer sets that clear (or approximately clear) the market. The principle of voluntary market exchange in and of itself is not sufficient to determine the market transaction. Thus, entropy equilibrium is a short-run, temporary equilibrium model of market exchange which replaces the Walrasian picture of the market in equilibrium as a budget hyper plane defined by equilibrium relative prices with a scalar field of transaction probabilities. (Brown & Shannon, 1997).

Under these conditions and following the same authors, entropy prices clear the market by distributing agents over their offer sets, rather than moving agents to optimal commodity bundles in their consumption sets, and thus effectively “convexify” the economy. Furthermore, the fact that different traders experience different transaction prices implies that random statistical equilibrium does not exhaust all the potential Pareto-improving transactions in the economy. Thus, such a statistical equilibrium approximates, but does not achieve, Pareto-efficiency. The statistical equilibrium in this market fails to achieve Pareto-efficiency because some potentially mutually advantageous transactions fail to be executed, and there is dispersion in actual transactions prices.

To achieve the Pareto optimum, the pioneering work of Foley (1994) has attempted to endogenize the offer sets of economic actors in a rational expectation framework:

> If we imagine a given agent repeatedly entering a market in statistical equilibrium, it is tempting to suppose that she will alter her offer set in order to optimize her market outcome given the probabilities that govern transactions in the market equilibrium. This idea gives rise to the concept of endogenous offer sets.

According to the same author, the present state of research should be that we rigorously establish that Walrasian equilibrium is the asymptotic outcome of a process in which endogenous offer sets adapt to statistical equilibrium entropy prices, and where the chances to transact in any period become numerous. Walrasian equilibrium is not unique, whereas statistical equilibrium for given offer sets of traders should be unique. The adaptation process sketched out here allows offer sets to change over time, giving rise to a dynamic process which may have multiple equilibria, each corresponding to a Walrasian equilibrium in markets with multiple Walrasian equilibria.

We argue that the problem should be placed in the context of non-additive statistics, suggesting that agent behaviours are time or space dependent. In fact, the above Gibbs-Shannon entropy related price suggests an ergodic system in which agent actions are disconnected from long-run memory of past market events and/or confluent—space related—information from surrounding markets. Then, system complexity describing market transactions should lead to non-extensive entropy related prices. Further research is needed to understand to what extent macroeconomic equilibrium and thermodynamic equilibrium are comparable in the context of non-ergodic real world systems.

To conclude, there is no assurance that the balanced post-entropy social accounting matrix is achieving (or approximating) the Pareto Optimum. As stated above, some probabilistic distributions of price entropy may be meaningless or not optimal in the convex space of all possible transactions—not even the voluntarily contracted ones.

More investigation is required to better appreciate, among other possibilities, the new approach of endogenizing dynamic offer sets, as suggested above. For the time being, entropy econometrics as presented in the coming sections seems to be the best approach to resolve such ill-behaved inverse problems.

3 The Social Accounting Matrix (SAM) Framework

3.1 Generalities

The objective of a social accounts matrix is to provide to economic analysts a highly detailed record, consistent with the relations between different economic agents at a particular moment in time, usually the period of a year, presenting in one matrix the interaction between production, income, consumption, and investment. A SAM is the representation of the circular flow established in the economy by the flow of money, on the one hand, and the flow of goods and services, on the other. A SAM constructed on a country or regional basis remains mainly static. Below we present a prototype of a SAM.

SAMs are square matrices (columns reflect transposed rows) in the sense that all economic sectors or institutional agents (Firms, Households, Government, and 'Rest of Economy' sector) are both buyers and sellers. Columns represent buyers (expenditures) and rows represent sellers (receipts). SAMs were created to identify all monetary flows from sources to recipients, within a disaggregated national account. The SAM is read from column to row, so each entry in the matrix comes from its column heading, going to the row heading. Each column total equals each corresponding row total. As already said, this ensures accounting consistency in the context of general equilibrium.

The initiator of the SAM table is the Cambridge Growth Project in Cambridge, England, which developed the first SAM in 1962 (Brown & Stone, 1962). This table was built as a matrix representation of the National Account. Finally, it was Graham (1985), the former associate of Stone in the 1960s who transferred SAM construction knowhow to the World Bank and became with E. Thorbecke, one of its leading proponents and developers on a worldwide scale.

One of the advantages of the SAM is it simplifies the design of the economy being modelled. As we will see in the next sections, a SAM forms the backbone of computable general equilibrium models or various types of empirical multiplier models.

The empirical importance of the SAM has influenced its form[30], which clearly depicts spending patterns of an economy. SAMs are currently in widespread use in the world, and many statistical bureaus, particularly in OECD countries, create both a national account and this matrix counterpart.

As in the case of input-output matrices, a theoretical SAM should always balance, but empirically estimated SAMs never do so in the first collation. The source of these imbalances may be diverse, but the principal ones are the conversion of national accounting data into money flows and the introduction of non-SNA. Additionally,

30 IMPLAN and RIMS II data exploit these standardized data formats to implement much economic impact analysis.

organizational problems in some countries can lead to conflicting sources of information, and the impossibility of balancing the constructed SAM creates a challenge. This problem was noted as early as 1984 by Mansur and Whalley (1984) and numerous techniques have been devised to 'adjust' SAMs, as "inconsistent data estimated with error." These techniques are similar to those discussed in the preceding chapter in the case of input-output tables, so they will not be discussed again here. As in the case of input-output matrices, when a SAM table is not balanced, this constitutes a case of the ill-behaved inverse problem. In fact, under this situation, besides a reduced number of accounts that we may claim to know with sufficient certainty, most of the accounts, about n(n – 1) of them, are known with uncertainty and the problem of indeterminacy usually arises. Robinson et al. (2001) first suggested a Shannon-Gibbs-Golan entropy based method for adjusting an unbalanced SAM of Mozambique's economy. Since then, as we will see later, a growing number publications applying this technique have been observed.

3.2 Description of a Standard SAM

The SAM table displays, in a systematic way, the circular flows inside macroeconomic accounts according to the bookkeeping principle (input equals output). Within that table, each account on the row stands for receipts by the sector at the left side of the same row. At the same time, these accounts constitute expenses by the sector situated on the top of the column. For instance, *sector activities* purchases *"Intermediate Inputs"* from the *commodity sector.* This means that this last sector receives a certain revenue from the *activity sector*.

Elements in the first column of *sector activity* (production sector) concern costs necessary to produce the global product at basic price. Thus, in this column are intermediary inputs (including imported goods) offered by different economic sectors, primary inputs (wages, operating surplus plus depreciation) and net indirect taxes (taxes on production minus subsidies from the government) on production. The second *sector of commodities* buys commodities from the *activity sector* on which direct taxes on consumption must be imposed — mainly value added taxes and import tariffs and duties. The commodity sector purchases goods from the rest of the world, i.e., *imports*, for internal consumption needs, complementary to commodities locally produced. Primary factors pay incomes from sold services to their owners (households, enterprises) in the activity sector. Primary factors disburse social security premiums to social institutions in exchange for insurance.

Households, thanks to income received from factors of production or public transfers, purchase goods for private consumption, pay direct taxes to the government, as well as taxes on wealth and other revenues and different non-tax disbursements in favour of local government. Households save income, not disbursed, for current needs and charges, and private banks receive these amounts for savings.

Enterprises disburse a part of their income in favour of their owners, i.e., households. Enterprises pay corporate taxes and the public sector factor in favour of government.

Next, social security institutions, which have received social security premiums from households, disburse social security expenditures in relation with insurance contract execution.

Government institutions disburse received income (mainly from taxes and other government charges) to consume commodities, to make transfers to households, and grant subsidies to enterprises, pay interest on domestic and foreign debts, and save the remaining income.

Domestic banks, like any other private enterprise, distribute net profits to households, pay taxes to the government, and disburse saved money for investment; they may also lend (national bank) money to the government to finance the difference between public investment and public savings. Finally, domestic banks pay interest to foreign banks.

Capital Account represents durable materials purchased from commodities for private or public investments.

Rest of the world (ROW), that is, foreign countries and institutions, import (our exports) goods and services locally produced, and pay remittances to national production factors. It pays private transfers of our enterprises residing in foreign countries. This account appears as a balance, taking into account transfers made by foreign enterprises residing inside our country. Rest of the world finances domestic banks, too. This happens, in particular, when branches of international banks are located inside the national territory.

In spite of the sectorial equilibrium between demand and supply predicted by theory—that is, in terms of the social accounting matrix, balancing respective rows and columns—the system of data collecting and organization of national accounting are far from perfect. As a consequence, we get a social accounting matrix which is unbalanced, at least for a few of the sectorial accounts.

This is generally the case for institution accounts like households, enterprise, government, and non-profit organizations. That is, contrary to the above explained conservation law, general equilibrium is not attained. For instance, total revenue of enterprises is lower than their total expenditures, and the situation is reversed in the case of the government. The principles of market clearance and of zero-profit may not be satisfied in the case of the enterprise sector. The principle of balanced budget account is not satisfied in the case of government. In practice, the case of no balanced accounts constitutes the rule rather than the exception. To balance respective rows and columns, economists may implement arbitrary economic hypotheses such as those related to closure rules[31] or use different quantitative techniques already dis-

31 The essential problem is that the classic CGE model, in which all markets clear, yields a full-emcussed above in the case of balancing input-output tables. Some of these techniques are more or less reliable. This is the case for the least squares or the linear programming methods, but their accuracy remains poor, as already noted. The bi-proportional RAS method (Golan et al., 1996) is relatively powerful. However, it is a not appropriate for the imposition of restrictions with random errors in the model.

Fortunately, recent techniques of cross-entropy econometrics display interesting estimation properties and seem to be adapted to such ratio problems. The next section presents the cross-entropy technique which will be applied to balance a social accounting matrix of Gabon (see output details in Bwanakare (2013)) or to present further applications of the procedure by balancing the ecological SAM of Poland (see Tables 23 and 24). As is well known, disaggregating a SAM requires a lot of, often non-reconcilable, information which worsens its equilibrium. This is particularly true in developing countries. The case of disaggregating labour or household accounts without reliable information to split them up should be emphasised.

ployment equilibrium and market-clearing prices, while short-run macro models typically involve wage and price rigidities, partial adjustment mechanisms, and equilibrium without market clearing, including unemployment. The two paradigms embody very different notions of equilibrium. http://www.ifpri.org/events/seminars/2003/20031014/robinson_thorbecke_EPIAM.pdf

4 Balancing a SAM

As already noted, the numerous data sources used in the process are in diagreement and in need of supplemental assumptions. Hence, it is no surprise that the resulting SAM is not balanced. The largest discrepancies are found inside accounts concerned with secondary distribution of income. This concerns, in particular, the disaggregated household account because of the assumptions that have been involved to spread the various incomes, transfers, and expenditures of households within its disaggregated components or between the disaggregated components and other accounts.

Once again, there are several ways of balancing inconsistent social accounting or any other matrix. Besides economic closure rules, one of the most commonly used techniques to balance matrices is the RAS approach. It is typically used for updating SAMs for which *new row* and *column sums* are known. As we have already noticed in the case of the input-output transactions matrix, the RAS technique produces a new transaction matrix that is consistent with the new row and column sums by interactively adjusting the row and column entries proportionately until the new totals are obtained. This approach has at least one severe drawback. The RAS technique assumes that the initial SAM is consistent and that there is no measurement error in the row and column sums. When dealing with social accounting matrices in general, the initial SAM will often not be consistent, there will typically be measurement errors, and there will certainly be some data entries that the analyst finds more reliable than others. Mainly for that reason, the cross-entropy approach seems to be better adapted for balancing the SAM. If a reader is interested in comparing the RAS approach with the cross-entropy technique, we refer him to Part III, section 3.3. The only question that remains concerns the distribution law of the model to be retained, capable of describing the above discrepancies and stochastic errors. This will be answered in the following section.

Since much has been said in Part II concerning relationships between the cross-entropy approach and the statistical information theory approach, only a concise presentation of this technique will be given here, and the reader is referred to the aforementioned references for further detail.

The entropy technique is a method of solving undetermined estimation problems. The problem is undetermined because, in the case of a SAM, for an *nxn* matrix dimension, we seek to identify n(n − 1) independent, unknown, non-negative parameters, i.e., all the cells of the SAM but one column or row, in conformity with Walras's law. In other words, restrictions must be imposed on the estimation problem so that we have enough information to obtain a unique solution and to provide enough degrees of freedom. The underlying philosophy of entropy estimation is to use all the information at hand for the problem and only that information: the estimation procedure should not ignore any available information nor should it add any false information.[32]

32 See Shannon (1948) and Theil (1967) for a discussion of the concept of 'information.'

In the case of a SAM estimation, 'information' may be the knowledge that there is measurement error concerning the variables, and that some parts of the SAM are known with more certainty than others. There may be a prior in the form of a SAM from a previous year, whereby the entropy problem is to estimate a new set of coefficients 'close' to the prior using new information to update it. Furthermore, 'information' could consist of moment constraints on, for instance, row and column sums, or the average of the column sums. In addition to the row and column sums, 'information' may also consist of certain economic aggregates such as total value-added, final demand components, and/or imports. In that way, it becomes possible[33] to maintain Walrasian conditions of equilibrium. Such information may be incorporated as linear adding-up restrictions on the relevant elements of the SAM. In addition to equality constraints such as these, information may also be incorporated in the form of inequality constraints to the macro-aggregates mentioned. In most cases, macroeconomic theory can be useful in suggesting signs or interval of variation of certain parameters or ratios. This information will then be incorporated among other constraining equations. Finally, one may want to restrict cells that are zero in the prior to remain so after the entropy balancing procedure. Similarly, some cell values belonging to the SAM to be updated may not need to be modified because they come from well documented sources. Such cell values could then be restricted to stay unmodified during all steps of information processing.

4.1 Shannon-Kullback-Leibler Cross-Entropy

Let us follow for the next estimation procedure found, for example, in Robinson et al. (2001), and let the SAM be defined as a matrix *T* with elements T_{ij} representing a payment from column account *j* to row account *i*. As mentioned above, each account is supposed to display Walrasian equilibrium. In other words, every row sum (tot_i) in the SAM must equal the corresponding column sum (tot_j):

$$tot_i = \sum_j T_{ij} = \sum_j T_{ji} \tag{4.1}$$

Dividing each cell entry in the matrix by its respective column total generates a matrix of column coefficients A:

$$A_{ij} = \frac{T_{ij}}{tot_j} \tag{4.2}$$

33 This constitutes a necessary and not a sufficient condition for the achieving of Walras equilibrium.

It is assumed that the entropy problem starts with a prior A which plausibly is a SAM from a previous period or, as in this case, a raw and unbalanced SAM. A represents the starting point from which the cross-entropy balancing procedure departs in deriving the new matrix of coefficients A. The entropy problem is to find a new set of A coefficients which minimize the so-called Kullback-Leibler (1951) divergence measure of the 'cross-entropy' (CE) between the prior A^* and the *posteriori* coefficients matrix A.

$$\min I(A \| A^*) = \sum_i \sum_j A_{ij} \ln \frac{A_{ij}}{A^*_{ij}} = \sum_i \sum_j A_{ij} \ln A_{ij} - \sum_i \sum_j A_{ij} \ln A^*_{ij} \tag{4.3}$$

subject to

$$\sum_j A_{ij} tot_i = tot_i \tag{4.4}$$

$$\sum_j A_{ij} = 1 \text{ and } 0 \leq A_{ij} \leq 1 \tag{4.5}$$

Note that, according to Walras's law in general equilibrium theory, one equation can be dropped in the second set of constraints: If all but one column and row sums are equal, the last one must also be equal. The solution of the above problem is obtained by setting up the Lagrangian. The k macro-aggregates can be added to the set of constraints on the problem above as follows:

$$\sum_i \sum_j \mathrm{H}_{ij}^{(k)} T_{ij} = \gamma^{(k)} \tag{4.6}$$

where H is an *nxn* aggregator matrix with ones for cells that represent the macro-constraints and zeros otherwise, and γ is the value of the aggregate constraint. As mentioned above, in the real world one faces economic data measured with error. The cross-entropy problem can also be formulated as an 'error-in-variables' system where the independent variables are measured with noise *e*. If, for example, we assume the known column sums are measured with error, the row/column consistency constraint can be written as:

$$tot_j = x_i + e_i \tag{4.7}$$

where tot_j is the vector of row sums and x_i, the known vector of column sums, is measured with error e_i. The prior estimate of the column sums could be, for instance, the initial column sums, the average of the initial column and row sums, or the row sums.

Following Golan et al. (1996), the errors are written as weighted averages of known constants *v* defined over a finite discrete support space $m>>1,\ldots,M$ with points:

$$e_i = \sum_{m>>1,\ldots,M} f_{im} v_{im} \tag{4.8}$$

where f_{im} is a set of weights that fulfil the following constraints:

$$\sum_{m>>1,\ldots,M} f_{im} = 1 \text{ and } 0 \le f_{im} \le 1 \tag{4.9}$$

In the estimation problem, the weights are thus treated as probabilities to be estimated, and the prior for the error distribution in this case is chosen to be a symmetric distribution around zero with predefined lower and upper bounds, and using either three or five weights. Naturally, not only the column and row sums can be measured with error, the macro-aggregates by which we constrain our estimation problem may also be measured with error, and so we can operate with two sets of errors with separate weights f_1's on the column sum errors, and weights f_2's on the macro-aggregate errors. The optimization problem in the 'errors-in-variables' formulation is now the problem of finding A's, f_1's, and f_2's that minimize the cross-entropy measure, including terms for the error weights:

$$\min I(A \| A^* \cup f_1 \| f_2) = \left(\sum_i \sum_j A_{ij} \ln A_{ij} - \sum_i \sum_j A_{ij} \ln A^*_{ij} \right) +$$

$$+ \left(\sum_i \sum_{m,\ldots,M} f_{1im} \ln f_{1im} - \sum_i \sum_{m,\ldots,M} f_{1im} \ln f_{1iJ} \right) + \left(\sum_i \sum_{m,\ldots,M} f_{2im} \ln f_{2im} - \sum_i \sum_{m,\ldots,M} f_{2im} \ln f_{2iJ} \right) \tag{4.10}$$

Referring once again to the definition of information provided by Kullback and presented in the second part of this book, cross-entropy measurements reflects how much the information we have introduced has moved our solution estimates away from the inconsistent prior, while also accounting for the imprecision of the moments assumed to be measured with error. Hence, if the information constraints are binding, the distance from the prior will increase. If they are not binding, the cross-entropy distance will be zero. It becomes now clearer why we have proposed, while assessing forecasting performance of the entropy technique, the difference between the average error variance coefficients (AEVC) of the periods 2006 and 2007 as a benchmark, maximum divergence precision measurement.

4.2 Balancing a SAM through Tsallis-Kullback-Leibler Cross-Entropy

In the following, we are going to generalize the Jaynes-Kullbac-Leibler model (4.10) and thus reconsider all the implications of the above theorem on the power law property of economy.

Let us formally explain Tsallis relative entropy model (2.47-2.50) to be minimized together with the above suggested constraints. In this presentation, the Bregman form of relative entropy (2.47) will be used:

$$\min I_q\left[A \| A^* \cup w \| wo\right] = \alpha \frac{1}{q-1} \sum_i A_{iJ}\left[(A_{iJ})^{q-1} - (A^*_{iJ})^{q-1}\right] - \sum_i (A_{iJ} - A^*_{iJ})(A^*_{iJ})^{q-1} +$$

$$(1-\alpha)\frac{1}{q-1}\sum_h w_{ih}\left[(w_{ih})^{q-1} - (wo_{ih})^{q-1}\right] - \sum_h (w_{ih} - wo_{ih})(wo_{ih})^{q-1} \quad (4.11)$$

subject to:

$$\sum_j A_{ij} tot_j = tot_i \quad (4.12)$$

$$\sum\nolimits_{i>2\ldots M}^{N} A_{ij} = 1 \quad (4.13)$$

$$\sum\nolimits_{h>>1,\ldots,I} w_{ih} = 1 \quad (4.14)$$

Symbols are as in Equation (4.9), except w_{ih}, which takes the place of f_{2ij}, and both represent disturbance errors on parameters but, this time, of different distribution laws.

Empirical, long practice with this class of economy-wide models provides some prior information on relevant ranges for parameter values and likely parameter estimates. Furthermore, while the support of any imposed prior distribution for a parameter is a maintained hypothesis (the estimate must fall within the support), the shape of the prior distribution over that support (e.g., the weights on each support point) is not. Unless the prior is perfect, the data will push the estimated posterior distribution away from the prior. The direction and magnitude of these shifts are, in themselves, informative. Also, note from Equations (4.11– 4.14) that, with increases in the number of data points, the second term of prediction in the objective function increasingly dominates the first term *precision*. In the limit, the first term in the objective becomes irrelevant. The prior distributions on parameters are only relevant when information is scarce.

4.3 A SAM as a Generalized Input-Output System

In the present paragraph, Kullback-Leibler (K-L) information divergence is extended to Tsallis non-ergodic systems and a q-Generalization of the K-L relative entropy criterion function (c.f.), with *a priori* consistency constraints, is derived for balancing a SAM as a generalized input-output transaction matrix.

On the basis of an unbalanced, Gabonese social accounting matrix (SAM) representing a generalized inverse problem input-output system, we propose to update and balance it following the procedure explained through the above section.

4.3.1 A Generalized Linear Non-Extensive Entropy Econometric Model

This section applies the results of, e.g., Jaynes (1957) and Golan et al. (1996) to present the model to be later implemented for updating and balancing input-output systems. While the argument in the criterion function is already known (see Equation 4.18), we need to reparametrize[34] the generalized linear model, to be introduced later into the model as restrictions in the spirit of Bayesian method of moments (e.g., Zellner, 1991). Note that such a linear restriction will be affected by a stochastic term expected to belong to the larger family of power law distribution. Let us succinctly present the general procedure for parameter reparametrization as it follows:

$$Y = X \cdot \beta + \varepsilon \tag{4.15}$$

Parameter β in general bears values not constrained between 0 and 1. When this is the case, reparametrization will no longer be necessary since parameter variation area fits well to probability definition area. The variable ε is an unobservable disturbance term with finite variance, owing to the economic data nature of exhibiting observation errors from empirical measurement or from random shocks. These stochastic errors are assumed to be driven by a large class of PL. As in classical econometrics, variable Y represents the system, the image of which must be recovered, and X accounts for covariates generating the system with unobservable disturbance ε to be estimated through observable error components e. Unlike classical econometric models, no constraining hypothesis is needed. In particular, the number of parameters to be estimated may be higher than the observed data points and the quality of collected information data low. Additionally, as already explained, to increase the accuracy of such estimated parameters from the poor quality of data points, the entropy objective function allows for incorporation of all constraining functions which act as Bayesian *a priori* information in the model.

Let us treat each $\beta_k (k = 1,\ldots,K)$ as a discrete random variable with compact support and $2 < M < \infty$ possible outcomes. Thus, we can express β_k as:

$$B_k = \sum_{m=1}^{M} p_{km} v_{km} \quad \forall k \in K \tag{4.16}$$

where p_{km} is the probability of outcome v_{km} and the probabilities must be non-negative and sum up to one. Similarly, by treating each element e_i of e as a finite and discrete

34 Reparametrization aims at treating parameters of the model as outputs of probability distribution to be estimated following the procedure presented by Golan et al. (1996) and later exploited for modelling many entropy econometric models (see, e.g., Bwanakare et al. (2014, 2015, 2016). Since the same probabilities are related to entropy variable defining the criterion function, optimizing the whole model then leads to outputs taking into account stochastic *a priori* information owing to model restrictions.

random variable with compact support and $2 < M < \infty$ possible outcomes centred around zero, we can express e_i as:

$$e_i = \sum_{j>>1,\ldots,J} r_{nj} z_{nj} \tag{4.17}$$

where r_n is the probability of outcome z_n on the support space *j*. Following Bwanakare (2014), we will use the commonly adopted index *n*, here and in the remaining mathematical formulations, to set the number of statistical observations. Note that the term *e* can be initially fixed as a percentage of the explained or endogenous variable, as an *a priori* Bayesian hypothesis. Posterior probabilities within the support space may display non-Gaussian distribution. The element v_{km} constitutes an *a priori* information provided by the researcher while p_{km} is an unknown probability whose value must be determined by solving a maximum entropy problem. In matrix notation, let us rewrite $\beta = V \cdot P$

with $p_{km} \geq 0$ and $\sum_{k=1}^{K} \sum_{m>>1,\ldots,M} p_{km} = 1$,

where again, *K* is the number of parameters to be estimated and *M* the number of data points over the support space. Also, let $e = r \cdot w$

with $r_{nj} \geq 0$ $\sum_{n=1}^{N} \sum_{j>>1,\ldots,J}^{J} r_{nj} = 1$

and $r_{nj} = 1$ for *N* the number of observations and the *K* number of data points over the support space for the error term. Then, the Tsallis cross-entropy econometric estimator can be stated as:

$$MinH_q(p \| p^0, r \| r^0, w \| w^0) \equiv \alpha \sum p_{km} \frac{[p_{km} / po_{km}]^{q-1} - 1}{q-1} + \beta \sum r_{nj} \frac{[r_{nj} / ro_{nj}]^{q-1} - 1}{q-1} +$$

$$+ \delta \sum w_{ts} \frac{[w_{ts} / wo_{ts}]^{q-1} - 1}{q-1} \tag{4.18}$$

S. to

$$Y = X \cdot \beta + e = X \cdot \sum_{m=1}^{M} v_m \left(\frac{p_{km}{}^q}{\sum_{m=1}^{M} p_{km}{}^q} \right) + \sum_{j=1}^{J} z_j \left(\frac{r_{nj}{}^q}{\sum_{j=1}^{J} r_{nj}{}^q} \right) \tag{4.19}$$

$$\sum_{k=1}^{K} \sum_{m>2,\ldots,M} p_{km} = 1$$

$$\sum_{n=1}^{N} \sum_{j>2,\ldots,J}^{J} r_{nj} = 1$$

$$\sum_{t=1}^{T}\sum_{s>2,\ldots,S}^{S} w_{ts} = 1 \tag{4.20}$$

Additionally, k macro-aggregates can be added to the set of constraints as follows:

$$\sum_{i}\sum_{j}\mathrm{H}^{(d)}{}_{ij}T_{ij} = \gamma^{(d)} + \sum_{s=1}^{S} g_s\left(\frac{w_{ts}{}^{q}}{\sum_{t=1}^{T} w_{ts}{}^{q}}\right), \tag{4.21}$$

where H is a dxd aggregator matrix with ones for cells that represent the macro-constraints and zeros otherwise, and γ is the expected value of the aggregate constraint. Once again, g_s stands for a discrete point support space from s = 2,…,s. Probabilities w_{ts} stand for point weights over g_s. The real q, as previously stated, stands for the Tsallis parameter.

Above, $H_q(p||p^0, r||r^0, w||w^0)$ is nonlinear and measures the entropy in the model. Relative entropies of three independent systems (three posteriors p, r, and w and corresponding priors p^0, r^0, and w^0) are then summed up using weights $\alpha\beta\delta$. These are positive reals summing up to unity under the given restrictions. We need to find the minimum divergence between the priors and the posteriors while the imposed stochastic restrictions and normalization conditions must be fulfilled. As will be the case in the application below, the first component of the criterion function may concern the parameter structure of the table; the second component errors on column (or row) totals and the last component may concern errors around any additional consistency variable, such as the GDP in the case below. As it has been shown by Tsallis (2009), this form of entropy displays the same basic properties as K-L information divergence index or relative entropy. The estimates of the parameters and residual are sensitive to the length and position of support intervals of β parameters (Equations 4.16 and 4.17) in the context of the Bayesian prior. When parameters of the proposed model are expressed under the form of elasticity or ratio—as will be the case in the example below—then the support space should be defined inside the interval between zero and one and will fit that of the usual probability variation interval. In such a case, no reparametrization of the model is needed. In general, support space will be defined between minus and plus infinity, according to the prior belief about the parameter area variation by the modeller. Additionally, within the same support space, the model estimates and their variances should be affected by the support space scaling effect, i.e., the number of affected point values (Foley, 1994). The higher the number of these points, the better the prior information about the system. The weights $\alpha\beta\delta$ are introduced into the above dual objective function. The first term of "precision" accounts for deviations of the estimated parameters from the prior (generally defined under a support space). The second and the third terms of "prediction ex-post" account for the empirical error term as a difference between predicted and observed data values of the model. As expected, the presented entropy model is an efficient information processing rule that transforms, according to Bayes's rule, prior and sample information into posterior information (Ashok, 1979).

4.4 Input-Output Power Law (Pl) Structure

It is time now to come back to the fundamental problem concerning the true statistical nature of input-output data used in the above studies or those below. In recent years, as already explained in Part I, many studies (Champernowne, 1953; Gabaix, 2008) have shown that a large array of economic laws take the form of a PL, in particular macroeconomic scaling laws, distribution of income and wealth, size of cities, firms[35], and the distribution of financial variables, such as returns and trading volume. Stanley and Mantegna (2007) have studied the dynamics of a general system composed of interacting units each with a complex internal structure comprising many subunits where the latter grow in a multiplicative way over a period of twenty years. They found the system follows a PL distribution. Such outputs should present similarities with the internal mechanism of national accounts tables, such as an input output table or a SAM. A PL displays, besides its well-known scaling law, a set of interesting characterizations related to aggregative properties of a PL according to which a power law is conserved under addition, multiplication, polynomial transformation, and minimum and maximum. As far as the PL hypothesis for a SAM is concerned, taking into consideration the above literature and using PL properties, it should not be difficult to prove the PL character of a SAM, including the Gaussian trivial case. About SAM construction and components, see for example, Pyatt (1985). General equilibrium (Wing Ian Sue, Sept 2004) implies that respective row and column totals are expected to balance. Conceptually, this model is based on the laws of product and value conservation (Serban Scrieciu & Blake, 2005) which guarantee conditions of zero profits, market clearance, and income balance. However, different stages of statistical data processing remain concomitant with human errors and the SAM will not balance. It is generally assumed that the main sources of these imbalances remain different sources of documentation and different time of data collecting. This means that an unknown number of economic transaction values within the matrix are inconsistent with the data generating macroeconomic system. For clarity, let us use Table 13 to explain these imbalances, noting, for instance, a difference between the institution row and column totals as follows:

$$(\mathrm{iT} + e_4) - (\mathrm{iT} + \varepsilon_4) = (e_4 - \varepsilon_4) \tag{4.22}$$

35 See Bottazzi et al. (2007) for a different standpoint on the subject.

Table 11: General structure of a stochastic non-balanced SAM

	Activities	Commodities	Factors	Institutions	Capital	World	Total
Activities	0	Ac	0	Ai	0	aw	aT+ ε_1
Commodities	Ca	0	0	Ci	cc	0	cT+ ε_2
Factors	Fa	0	0	0	0	0	fT+ ε_3
Institutions	Ia	Ic	If	ii	0	iw	iT+ ε_4
Capital	0	0	0	ci	0	cw	cT+ ε_5
World	0	Wc	0	wi	0	0	wT+ ε_6
Total	aT+ e_1	cT+ e_2	fT+ e_3	iT+ e_4	cT+ e_5	wT+ e_6	

Source: own elaboration.

The term on the left hand side of the above expression represents the difference between two erroneous and unequal totals of *institution* account. The origin of that difference results from difference between plausibly different stochastic errors e_4 and ε_4, respectively, on column and row totals. In Table 11, the first alphabetical letter of symbols inside each cell represents the first letter of the row (supply) account, and the second letter represents the first letter of the corresponding (demand) column. In the SAM prototype below, e.g., the symbol "Ca", explains purchases by the activity sector of goods and services from the commodity sector.

The targeted purpose is to find, out of all probability distributions, a set of *a posteriori* probabilities closest to *a priori* initial probabilities and insure the balance of the SAM table while satisfying other imposed consistency moments and normalization conditions. Following Shannon terminology, one may consider post-entropy structural coefficients and disturbance errors, respectively, as signal and noise. The first step consists in computing *a priori* coefficients by column from real data from Table 11 by dividing each cell account by the respective column total. Next, we treat these column coefficients as analogous to probabilities and column totals as expected column sums, weighted by these probabilities (see Equation 4.19). These coefficient values will serve as the starting, best prior estimates of the model. The other two types of priors to initialize the solution concern errors on column totals (Equation 4.17) and on gross domestic product (GDP) at factor and market prices (Equation 4.21). GDP variables are added to the model with the purpose of binding the latter to meet consistency macroeconomic relationships for different accounts inside the SAM. Other macroeconomic relations like those affecting interior or global consumptions could be added. The proposed approach combines non-ergodic Tsallis entropy with Bayes' rule to solve a generalized random inverse problem. We may optionally consider only some cell values as certain[36] while the rest of the accounts are unknown. This is one of the strongest points of the entropy approach over other mechanical techniques of balancing the national accounts table through a stochastic framework. All row and column totals are known with uncertainty. It is straightforward to notice that the potential freedom degree number of parameters to estimate $(n-1)(n-1)$ remains significantly higher than n observed data points (column totals). In a particular case of a SAM, and due to empty cells, that number of unknown parameters may be much lower. In any event, that will not generally prevent us from dealing with an ill-behaved inverse stochastic problem. The next important step is initializing the above defined errors through a reparametrizing process. *A five point support space symmetric around zero* is defined. To scale the error support space to real data, we apply Chebychev's inequality and three sigma rules (Serban Scrieciu & Blake, 2005). Corresponding optimal probability weights are then computed so as to define the prior noise component (Robinson et al., 2001).

4.5 Balancing a SAM of a Developing Country: the Case of the Republic of Gabon

In our analysis of the last cases, we have rather underscored technical aspects of entropy for balancing input-output tables. However, when statistical data from different sources are available and sufficiently consistent, applying a complex procedure as the one relying on entropy formalism can be more time consuming than relatively easier techniques like the RAS (e.g., Pukelsheim, 1994; Bacharach, 1970). This is the case for many developed countries where statistical data gathering is generally efficient[37]. On the contrary, as we are going to see in the coming pages, this is not the case for the majority of developing countries in which statistical data are not only scarce but also of bad quality.

Thus, to complete an array of empirical advantages of the proposed entropy approach, we are going to analyse the case of developing countries where complete

36 Only transaction accounts with the rest of the world (import, export, external current balance), plus government commodity consumption accounts are concerned.

37 The statistical data gathering system in Poland can been seen as relatively efficient in comparison with those of most of developing countries. Availability of data on a large scale and their quasi-consistency though from various unrelated sources is the criterion retained here for giving such an appraisal. As a result, it should be relatively easier to balance national account tables without using complicated procedures such the entropy-related one. Thus, inconsistencies displayed in Table 1 may not reflect outputs from other publications on the same subject. The purpose of the present example is just to show the performance of the cross-entropy procedure in balancing a system under constraining, *a priori* information, like different macroeconomic identities characterizing national account tables.

statistical information is generally unavailable. Not only does such information not fully exist, what does should be approached with a high level of uncertainty. Applying traditional balancing techniques, like the RAS approach, becomes in practice difficult.

Based on the Shannon entropy approach, a large number of studies—particularly from developing countries—designed to balance SAM tables have been prepared in the last two decades. The already cited paper of Robinson et al. (2001), consecutive to the publications of Golan et al. (1996), has become a reference work for having shown an algorithm—in GAMS code (General Algebraic Modelling System)—for balancing a SAM in the case of uncertainty. One can list other studies with identical purpose, such as those of Salem (2004) for Tunisia and Kerwat et al. (2009) for Libya. Murat (2005), using Shannon cross-entropy formalism, has balanced a Turkish Financial Social Accounting Matrix and, more recently, Miller et al. (2011) has built and balanced a disaggregated SAM for Ireland. Note that these last two countries belong, respectively, to the category of intermediary developed and developed countries. Many other entropy-based studies have been presented for various countries like Malawi, South Africa, Zimbabwe, Ghana, Gabon, and Vietnam. The results shown below generalize, once again, Shannon formalism by applying a non-extensive entropy divergence formalism.

4.5.1 Balancing the SAM of Gabon by Tsallis Cross-Entropy Formalism

A complete description of data sources or others details concerning the methodology of building the aggregated and disaggregated SAM of Gabon can be found in Bwanakare (2013).[38] That methodology has been proposed by Robinson et al. (2001) for balancing the SAM of Mozambique. Briefly, it consists of two steps in building the final SAM. In the first step, an aggregate and unbalanced SAM is built on the basis of official macroeconomic data. The second will serve as a control in building a much more disaggregated SAM in which accounts will be obtained by splitting out aggregated accounts of the balanced[39] SAM of the first step. Table 12 below represents the initial aggregated and unbalanced SAM of Gabon. Statistical data come from three sources: the Ministry of Planning, the Ministry of Economy and Finance, and the Bank of Central Africa States.

38 This document was prepared with the help of the Directorate of National Accounting at the Ministry of Planning and Development of the Republic of Gabon. A copy of the outputs of the balancing of this SAM has been transmitted to the Ministry. See the document at http://www.numilog.com/236150/Methodologie-pour-la-balance-d-une-matrice-de-comptabilite-sociale-par-l-approche-econometrique-de-l-entropie---le-cas-du-Gabon (ebook: Paris: Editions JePublie).

39 We have used the cross-entropy technique for balancing such an aggregated SAM.

Table 12: Aggregated non-balanced basic SAM of Gabon

	Activity	Product	Factors	Agents	enterprises	government	ROW	Accumulation	Total
Activities	-	4 784 615	-	103 019	-	-	-	-	4 887 634
Product	1 984 702	0	-	969 422	-	666 164	1 879 676	698 293	6 198 256
Factors	2 618 186	-	-	-	-	-	-	-	2 618 186
Agents	-	-	1 291 908	-	47 159	133 781	11 006	-	1 483 854
Enterprises	-	-	1 326 100	6 018	38 994	49 224	-	-	1 420 337
Government	284 746	215 181	1 029	104 043	484 284	742 353	9 621	-	1 841 256
ROW	-	1 198 460	- 850	201 885	241 948	145 179	-	-	1 786 623
Accumulation	-	-	-	99 467	559 791	152 715	- 113 680	-	698 293
Total	4 887 634	6 198 256	2 618 186	1 483 854	1 372 177	1 889 417	1 786 622	698 293	20 934 438

Sources: own calculations on the basis of official statistics of Gabon. (millions of CFA francs)

Thus, one may observe that all columns are not balanced with respective rows. This is the case for *the enterprise* and *government* accounts. This means that the general equilibrium is not attained. The total revenue of enterprises is higher than their total expenditures and the situation is reverse in the case of government.

In the case of Gabon, we target disaggregate accounts from 8 macro-accounts of the aggregated SAM to 82 subaccounts of the disaggregated new SAM. Among these subaccounts, we have 31 activity accounts, 33 product accounts (including 3 margin accounts which replace wholesale and retail trade), 3 accounts of labour factors and 1 of social capital, 4 classes of households, 3 business institutions, 4 government accounts, including public investment. The rest of the accounts are private investment, change in stock and rest of the world. Disaggregating a SAM requires much and often non-consistent information from different sources and periods. In the case of Gabon, this is particularly true for labour or the household accounts since reliable information allowing to tease out such accounts is scarce.

In statistical theory of information terms, the problem to be solved is one of finding a new disaggregated, *a posteriori* balanced SAM as close as possible to the initial unbalanced and disaggregated *a priori* SAM, while fulfilling imposed statistical and/or macroeconomic restrictions. To implement the model, we use the mathematical expression of non-extensive relative entropy under the next additional macroeconomic restrictions related to the targeted period (i.e., 1997 in the case of Gabon):

- Nominal GDP = consumption + investment + government expenditures + export – import
- GDP at factor cost = Nominal GDP – indirect taxes + subsidies
- Nominal GNP = Nominal GDP + net foreign income
- Fixing the input-output coefficients inside the SAM to the level of the previous period, implying that the structure of the Gabonese economy has not changed during the preceding years. Such a hypothesis remains realistic in the case of most developing countries, over a relatively long period.
- All accounts concerning business with the rest of the world have been fixed to the known level from international sources. This is so because, generally, data on international business remain reliable even in the case of developing countries.

When analysing the discrepancy between the prior and the posterior SAM, important modifications are observed. In particular, important discrepancies take place in the case of the institution and factor accounts. For instance, we note that wage assessment in the petroleum sector is probably underestimated for the senior executive category by around 300% of the real value deriving from post entropy modelling. In the period 1977, it had been pointed out many times by international institutions and media that financial transparency in Gabon needed to be improved. We note large modifications in factor inputs for bank and insurance activities.

Finally, it is important to note that outputs from Shannon cross-entropy (reported in Bwanakare, 2013) are identical to those from Tsallis cross-entropy formalism. As pointed out many times, this suggests that we are dealing with a Gaussian distribution model. Thus, since similar outputs have been published in the above reference, outputs from the non-extensive entropy technique are not presented in this monograph. Nevertheless, interested readers can obtain more details on these outputs.

To conclude, the question of assessing the performance of the approach could be posed here. In fact, since no previous, benchmark SAM exists in the case of Gabon[40], it is difficult to know to what extent we have deviated from values representing the true level of the economy. Fortunately enough, the entropy approach allows additional information embodied by the macroeconomic restrictions to be easily incorporated into the model. Next, when the optimum solution is reached, we then get the best results, generally conforming to our expectations. This should be the case in the present Gabon model.

4.6 About the Extended SAM

A SAM can be extended in different ways and for different purposes. Generally a SAM is extended to incorporate monetary aspects of the economy or to take into account the natural environment. In this document, we will deal with this last case only.

A SAM can be extended and incorporate auxiliary accounts concerning the environment and natural resource sectors, so that it becomes possible to analyse interactions between them and the economy. In fact, an environmentally extended SAM (ESAM) usually captures the relationships among economic activities, pollution abatement activities, and pollution emissions. The multiplier and structural path analyses are applied to the ESAM for assessing environmental impacts of pollution-related economic policies.

Recent literature shows that an ESAM can be a useful tool for environmental policy analysis. Interested readers can find rich and detailed information in the monograph of Plich (2002).

Table 13 presents a Polish unbalanced ESAM. The particularity of that matrix is that we have added four new sectors related to ecological activity. The first sector is the abatement ecological activity sector. In that sector, firms carry out depollution activities. The second sector is the abatement ecological commodity sector, which offers the produced services and products to the market. The third sector is that of pollution fees. Firms (or households) must pay tradable pollution permits or other forms of tax to government as a cost of using the polluting engines. In this context, this sector is considered as a (negative) factor of production Plich (2002). The last sector concerns the environmental capital accumulation for depolluting activity. It is

40 As in many developing countries, even if it existed, it would not necessarily represent a good reference.

worth emphasising that Table 13 may display more or less weakness since any benchmark table for Poland has not been found in the existing literature. The sole source of data used is the Polish Central Office of Statistics (GUS). Under these circumstances, let us suppose that more reliable information does not exist, as it often happens for this kind of research. Finding results using traditional approaches may take time, and reliability of outputs is generally limited since environmental data assessment is a difficult task. The purpose of this section is to apply entropy formalism to update the unbalanced Table 13. As we already know, the more significant moment restrictions are, the less significant precision errors will result. There are many reasons to consider the non-extensive entropy model to be—in this case of an ecological model— an ideal *balancing rule* since it has been proven to display multidisciplinary properties in many application areas. After having applied non-extensive cross entropy formalism (see Equations 4.18–4.21), we present below Table 14 an environmentally extended (aggregated) balanced Polish SAM (2005). In this experiment, accounts related to government and to foreign operations are supposed to be known with a random error. Such restrictions seem acceptable for a country like Poland, where statistical data on government incomes and expenses or operations with the rest of the world remain sufficiently reliable. The next restriction has concerned matrix cells with zero values in the initial Table 13. These zero value cells have been supposed to be known with certainty so that no change has modified them after computation. Information divergence between the two tables is reported in Table 15. Values inside the cells are in percent. The reader may notice the precision of the model, through the retained constraining variables in the minimization entropy model.

Table 13: An environmentally extended (aggregated) *unbalanced* Polish SAM (2005).

	aACT	aPOLLA-BAT	pCOM	PPOLLA-BAT	Labor	Capital	POLLFEES	Hou	Ent	GRE	CapAc	CapAcEnv	RoW	Total
aACT	0	0	19448.87	0.00	0.00	0.00	0.00	69.36	0.00	0.00	0.00	0.00	0.00	19518.23
aPOLLA-BAT	0	0	0	234.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	234.53
pCOM	10796.32	0.00	0.00	0.00	0.00	0.00	0.00	6989.91	0.00	784.48	1834.59	59.87	3646.58	24111.74
PPOLLA-BAT	65.41	0.49	0.00	0.00	0.00	0.00	0.00	167.70	0.00	0.00	0.00	0.00	0.00	233.60
Labor	3520.11	1.58	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3521.69
Capital	5048.40	5.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5053.61
POLLFEES	0.00	227.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	227.25
Hou	0.00	0.00	0.00	0.00	3503.16	2573.64	0.00	0.00	464.61	2711.41	0.00	0.00	375.28	9628.10
Ent	0.00	0.00	0.00	0.00	0.00	2126.84	0.00	0.00	0.00	42.46	0.00	0.00	77.89	2247.19
GRE	87.99	0.00	942.48	0.00	0.00	353.13	227.25	2241.24	0.00	353.13	0.00	0.00	58.63	4263.85
CapAc	0.00	0.00	0.00	0.00	0.00	0.00	0.00	652.48	1125.76	24.14	0.00	0.00	87.48	1889.86
CapAcEnv	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	29.93	29.93	0.00	0.00	0.00	59.87
RoW	0.00	0.00	3709.90	9.56	18.53	0.00	0.00	196.31	271.56	39.99	0.00	0.00	0.00	4245.85
Total	19518.23	234.53	24101.25	244.09	3521.69	5053.61	227.25	10317.00	1891.86	3985.55	1834.59	59.87	4245.85	

Source: own calculations on the basis of data from Polish Statistical Office (GUS) and from EUROSTAT.

Table 14:. An environmentally extended (aggregated) *balanced* Polish SAM (2005). (10 billion zł).

	aACT	aPOLLA-BAT	pCOM	pPOLLA-BAT	LABOR	CAPITAL	POLLFEES	HOU	ENT	GRE	CAPAC	CAPACENV	ROW	Total
aACT	0.00	0.00	19432.32	0.00	0.00	0.00	0.00	67.27	0.00	0.00	0.00	0.00	0.00	19499.60
aPOLLA-BAT	0.00	0.00	0.00	232.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	232.75
pCOM	10820.75	0.00	0.00	0.00	0.00	0.00	0.00	6785.67	0.00	818.21	1896.08	59.93	3654.35	24034.99
pPOLLA-BAT	71.47	0.48	0.00	0.00	0.00	0.00	0.00	170.09	0.00	0.00	0.00	0.00	0.00	242.04
LABOR	3543.13	1.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3544.68
CAPITAL	4977.47	5.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4982.59
POLLFEES	0.00	225.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	225.60
HOU	0.00	0.00	0.00	0.00	3526.05	2637.61	0.00	0.00	525.12	2827.23	0.00	0.00	375.97	9891.98
ENT	0.00	0.00	0.00	0.00	0.00	1987.15	0.00	0.00	0.00	41.00	0.00	0.00	72.12	2100.26
GRE	86.79	0.00	925.35	0.00	0.00	357.83	225.60	2126.06	0.00	364.76	0.00	0.00	58.18	4144.57
CAPAC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	553.12	1236.46	23.79	0.00	0.00	82.72	1896.08
CAPACENV	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	31.99	27.95	0.00	0.00	0.00	59.93
ROW	0.00	0.00	3677.31	9.29	18.63	0.00	0.00	189.77	306.70	41.64	0.00	0.00	0.00	4243.34
Total	19499.60	232.75	24034.99	242.04	3544.68	4982.59	225.60	9891.98	2100.26	4144.57	1896.08	59.93	4243.34	

Source: own calculations on the basis of data from Polish Statistical Office (GUS) and from EUROSTAT.

Table 15: Percentage of Information divergence between Tables 13 and 14.

	aACT	aPOLLA-BAT	pCOM	pPOLLA-BAT	LABOR	CAPITAL	POLLFEES	HOU	ENT	GRE	CAPAC	CAPACENV	ROW	Total
aACT	0	0.00	0.09	0.00	0.00	0.00	0.00	3.01	0.00	0.00	0.00	0.00	0.00	0.10
aPOLLABAT	0	0.00	0.00	0.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.76
COM	-0.23	0.00	0.00	0.00	0.00	0.00	0.00	2.92	0.00	-4.30	-3.35	-0.11	-0.21	0.32
pPOLLABAT	-9.27	1.63	0.00	0.00	0.00	0.00	0.00	-1.42	0.00	0.00	0.00	0.00	0.00	-3.61
LABOR	-0.65	1.77	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.65
CAPITAL	1.40	1.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.41
POLLFEES	0.00	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.73
HOU	0.00	0.00	0.00	0.00	-0.65	-2.49	0.00	0.00	-13.02	-4.27	0.00	0.00	-0.19	-2.74
ENT	0.00	0.00	0.00	0.00	0.00	6.57	0.00	0.00	0.00	3.45	0.00	0.00	7.40	6.54
GRE	1.36	0.00	1.82	0.00	0.00	-1.33	0.73	5.14	0.00	-3.29	0.00	0.00	0.77	2.80
CAPAC	0.00	0.00	0.00	0.00	0.00	0.00	0.00	15.23	-9.83	1.46	0.00	0.00	5.45	-0.33
CAPACENV	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-6.87	6.64	0.00	0.00	0.00	-0.11
ROW	0.00	0.00	0.88	2.82	-0.53	0.00	0.00	3.33	-12.94	-4.12	0.00	0.00	0.00	0.06
Total	0.10	0.76	0.27	0.84	-0.65	1.41	0.73	4.12	-11.02	-3.99	-3.35	-0.11	0.06	

Source: own calculations.

Glossary of table abbreviations:

aAct: activity sector
aPOLLABAT: abatement ecological activity sector
pCom: commodity sector
pPOLLABAT: abatement ecological commodity sector
Labor: labor sector (factor of production)
Capital: capital sector (factor of production)
Pollfees: pollution fees sector (factor of production)
Hou: households institution
Ent: enterprise institution
GRE: government institution
CapAc: capital accumulation sector (private investment)
CAPACENV: abatement -oriented capital accumulation sector (private investment)
RoW: rest of the world institution

5 A SAM and Multiplier Analysis: Economic Linkages and Multiplier Effects

5.1 What are the Economic Linkages and Multiplier Effects?

The strongest argument in favour of the Walras equilibrium—as opposed to the Marshall *ceteris paribus* approach—will find its momentum once industry linkages and multiplier effects are envisaged. This is so because in these circumstances thinking about partial equilibrium becomes less sustainable. In fact, the effect of a shock from one industry may have direct and indirect impact on the whole system defined by different industries. Let us analyse below a shock generated by the demand side. When we talk of "exogenous demand-side shocks" to an economy, we refer to changes to final control demand aggregates, i.e., export demand, government spending, or net investment demand of stocks. The effects of these shocks are both direct and indirect.

The direct effects are to those sectors that affront the shock. For example, an exogenous increase in demand for Polish manufactured exports has a direct impact on the manufacturing industry, which results in increased inputs, production, sales, and value-added. However, the positive consequences of such a shock go beyond the manufacturing industry. It may also have indirect effects stemming from manufactures' linkages to other industries inside the economy. These indirect linkages can be classified into supply-side and demand-side. When we add up all direct and indirect linkages, we get a measure of the shock's multiplier effect, or how much an initial effect is amplified or multiplied by indirect linkage effects. Supply-side linkages are determined by industry production technologies, which can be depicted from an input-output table. Next, they are differentiated into backward and forward linkages. Backward production linkages are the demand for additional inputs used by producers to supply additional goods or services. For instance, when production (of manufacturers) expands, it requires additional intermediate goods or services like raw material, machinery, and transport services. This demand then stimulates production of other industries that supply these intermediate goods. Technical coefficients supply information on the input intensity of the production technology used. The more an industry's production technology is input intensive, the stronger its backward linkages.

Forward linkages allude to supply inputs to upstream industries. For instance, increased manufacturer production should lead to increased supply of goods to the construction industry, which, in turn, stimulates, among others, service industries. As in the case of backward linkages, the more important an industry is regarding upstream industries, the stronger its forward linkages will be and multipliers will definitely become larger.

The conceptual structure of the input-output matrix only allows for deriving multipliers that measure the effects of supply linkages. Since the input-output table does not show secondary income distribution, it is not possible to consider consumption linkages, which arise when an expansion of production generates additional incomes for factors and households, which are then used to purchase goods and services. Continuing the same example as above, when manufacturing production expands, it raises households' incomes, which are used to buy consumer goods. Depending on the share of domestically produced, tradable, and imported goods in households' consumption baskets, domestic producers benefit from greater demand for their products. The size of consumption linkages depends on various factors, including the share of net factor income distributed to households; for an open economy, the level of gross domestic product per inhabitant, which exercises an influence on the composition of the consumption basket; and the relative price between locally produced and imported goods which determines in Armington fashion the share of domestically supplied goods in consumer demand.

Consequently, SAM multipliers tend to be larger than input-output multipliers because they capture both production and consumption/income linkages.

Following Breisinger et al. (2009),

> while economic linkages are determined by the structural characteristics of an economy (evidenced through technical coefficients and/or the composition of households' consumption baskets) and remain thus static, multiplier effects capture the combined dynamic effects of economic linkages over a period of time through different auto-generated rounds.

Three types of multipliers are generally reported in empirical research. First, an output multiplier combines all direct and indirect (consumption and production) effects across multiple rounds and reports the final increase in gross output of all production activities. Second, a GDP multiplier measures the total change value-added or factor incomes caused by direct and indirect effects. Finally, the income multiplier measures the total change in household incomes.

The dampening path of multipliers is consecutive to the level of leakage inside economic circular flows. Ultimately, higher leakages stemming from income allocated to imported goods or from government taxes make the round-by-round effects slow down more quickly and reduce the total multiplier effect.

In empirical research, one must often deal with two kinds of economic hypotheses. First, still in the context of the above example, one can suppose that demand shock will encounter no constrained response from the supply side. The second case is the one where demand shock is constrained. This can happen when supply is not able to completely satisfy increased demand. In this hypothesis, multipliers will follow a modified dynamic path towards slowing down. Let us still follow Breisinger et al. (2009) and then succinctly analyse both cases.

5.1.1 A SAM Unconstrained Multiplier

Let us present below a simplified SAM where presented accounts are just those required to derive a multiplier matrix.

Table 16: A simplified SAM for multiplier analysis

	Activities		**Commodities**		**Factors**	**Households**	**Exogenous demand**	**Total**
	A1	A2	C1	C2	F	H	E	
A1			X_1					X_1
A2				X_2				X_2
C1	Z_{11}	Z_{12}				C_1	E_1	Z_1
C2	Z_{22}	Z_{22}				C_2	E_2	Z_2
F	V_1	V_2						V
H					$V_1 + V_2$			Y
E			L_1	L2		S		E
Total	X_1	X_2	Z_1	Z_2	V	Y	E	

Source: own elaboration, based on Breisinger, Thomas, and Thurlow (2009).

We divide columns by their total to derive the coefficient matrix (M-matrix) excluding the exogenous components of demand.

Table 17: Transformed Table 16

	Activity		**Commodities**		**Factors**	**Households**	**Exogenous demand**
	A1	A2	C1	C2	F	H	E
A1			$b_1=X_1/Z_1$				
A2				$b_2=X_2/Z_2$			
C1	$a_{11}=Z_{11}/X_1$	$a_{12}=Z_{12}/X_2$				$c_1=C_1/Y$	E_1
C2	$a_{21}=Z_{21}/X_1$	$a_{22}=Z_{22}/X_2$				$c_2=C_2/Y$	E_2
F	$v_1=V_1/X_1$	$v_2=V_2/X_2$					
H					$1=(V_1 + V_2)/V$		
E			$l_1=L_1/Z_1$	$l_2=L_2/Z_2$		$s=S/Y$	
Total	1	1	1	1	1	1	E

Source: own elaboration, based on Breisinger, Thomas, and Thurlow (2009).

Symbols:

a) Values:
X: Gross output of each activity (i.e., X_1 and X_2)
Z: Total demand for each commodity (i.e., Z_1 and Z_2)
V: Total factor income
Y: Total household income
E: Exogenous components of demand shares

b) Share:
a: Technical coefficients
b: Share of domestic output in total demand
v: Share of value-added or factor income in gross output
l: Share of the value of total demand from imports or commodity taxes
c: Household consumption expenditure shares
s: Household savings rate

To derive equations representing the relationships in the above SAM, we start by setting up simple demand equations:

$$Z_1 = a_{11}X_1 + a_{12}X_2 + c_1Y + E_1$$

$$Z_2 = a_{21}X_1 + a_{22}X_2 + c_2Y + E_2 \quad (4.23)$$

Total demand = intermediate demand + household demand + exogenous demand.

The next relationships tell us that domestic production X is only part of total demand Z.

$$X_1 = b_1Z_1$$

$$X_2 = b_2Z_2$$

Since household income Y depends on the share each factor earns in each sector, then:

$$Y = v_1X_1 + v_2X_2$$

or,

$$Y = v_1b_1Z_1 + v_2b_2Z_2$$

Now replacing all X and Y in Equation (4.23), moving everything except for E onto the left-hand side, and grouping Z together, we finally obtain:

$$(I - M)Z = E, \quad (4.24)$$

where

$$(I-M)=\begin{pmatrix}(1-a_{11}b_1-c_1v_1b_1),(-a_{12}b_2-c_1v_2b_2)\\(-a_{21}b_1-c_2v_1b_1),\ (1-a_{22}b_2-c_2v_2b_2)\end{pmatrix}.$$

We note that M is a square matrix, the elements (share values) of which are not negative. Each column sum (see Table 16) is less or equal to unity. Thus, an inverse matrix of $(I - M)$ exists and should display non-negative values, suggesting the non-negativity of the multiplier matrix.

Formally, from (4.24) we directly get the final multiplier equation of the form:

$$Z = (I - M)^{-1}E \tag{4.25}$$

Total demand = multiplier matrix × exogenous demand

The above formulation tells us that when exogenous demand E increases, one will end up with a final increase in total demand equal to Z, owing to all the direct and indirect multiplier effects $(I - M)^{-1}$.

5.1.2 Equation System for Constrained SAM Multiplier

Often when factor allocation is not optimal, exogenous demand shocks may encounter limited response from producing sectors. Let us analyse below how much a multiplier will change if some producing sectors are unable to correctly respond. The expected issue is that if we fix one of two sectors Z, for instance Z_2. In that case, imports should substitute for domestic supply, thus eliminating any growth linkages from this sector.

The next equation is related to the non-constrained case and expresses total demand as the sum of its parts.

$$(1 - a_{11}b_1 - c_1v_1b_1)Z_1 + (- a_{12}b_2 - c_1v_2b_2)Z_2 = E_1$$

$$(- a_{21}b_1 - c_2v_1b_1)Z_1 + (1 - a_{22}b_2 - c_2v_2b_2)Z_2 = E_2$$

Grouping exogenous terms on the right-hand side (i.e., E_1 and Z_2) and rearranging[41], we finally obtain:

$$\begin{pmatrix}Z_1\\E_2\end{pmatrix}=\left(I-M^*\right)^{-1}B\begin{pmatrix}E_1\\Z_2\end{pmatrix} \tag{4.26}$$

where

$$\left(I-M^*\right)=\begin{pmatrix}(1-a_{11}b_1-c_1v_1b_1,0)\\(-a_{21}b_1-c_2v_1b_1,-1)\end{pmatrix}$$

41 For derivation details, see Breisinger, Thomas, and Thurlow (2009).

and

$$B = \begin{pmatrix} 1, a_{12}b_2 + c_1 v_2 b_2 \\ 0, -1 + a_{22}b_2 + c_2 v_2 b_2 \end{pmatrix}$$

Interpretation of the above equation is the following: an exogenous increase in demand for the unconstrained sectors $[E_1]$ leads to final increase in total demand for these sectors $[Z_1]$, including all of the forward and backward linkages $(I - M^*)^{-1}$. For the sectors with constrained supply (in our case sector Z_2), it is net exports that decline. This means that the current trade balance must worsen if we have to amortize demand shock in the case of constrained supply. If exports remained unchanged, then the alternative of reducing exports would be increasing imports so as to meet additional exogenous demand in the context of this constrained supply.

5.1.3 On Modelling Multiplier Impact for an Ill-Behaved SAM

Let us now return back to the central problem of this presentation and suppose that the matrix is unbalanced, which implies that multiplier values are not reliable. The way to avoid this should consist of only estimating parameters of the model without taking into account the obligation that the whole SAM be internally consistent. Thus, we should maximize (or minimize) entropy for probabilities related to the multiplier matrix under traditional restrictions, plus an additional constraint declaring values of an already balanced SAM to be taken as a prior.

Remembering about the interpretation of estimated parameters through the maximum entropy principle, it would be easy to make a link between the multiplier effect and maximum entropy modelling. In fact, in a linear model, parameters estimated by entropy formalism are interpreted as the long-run (equilibrium) impact of one unit change of regressor x on regresand y. Thus, long-run impact means that direct and indirect effects of the multiplier are accounted for with respect to the shock.

Annex C. Proof of Economy Power Law Properties

1. Definition of Power Law Distribution

Since we already know existing relationships between power law function and non-extensive entropy from Part II of this work, let us now present the main properties of the former in the context of a SAM.

Using a simplified formulation, a power law is the relation of the form $f(x) = Kx^{\alpha}$ where $x \succ 0$ and K and α are constants. While power laws can appear in many different contexts, the most common are those where $f(x)$ describes a distribution of random variables or the autocorrelation function of a random process.

The formulation above has the advantage of being intuitive. However, it does not show the real attributes of that distribution, which displays asymptotical characteristics.

Thus, the notion of a power law as it is used in extreme value theory is an asymptotic scaling relation. Let us first explain what we understand by equivalent scaling. Two functions f and g have equivalent scaling, $f(x) \sim g(x)$ in the limit $x \to \infty$[42] if:

$$\lim_{x \to \infty} \frac{L(x) f(x)}{g(x)} = 1 \tag{4.27}$$

with $L(x)$ is a slowly varying function, thus satisfying the relation:

$$\lim_{x \to \infty} \frac{L(tx)}{L(x)} = 1,$$

for any finite constant t > 0. Slowly varying functions are, for example, $L(x) = C$ and $L(x) = \ln(x)$, that is, a constant and a logarithmic function, respectively.

A power law is defined as any function satisfying $f(x) \sim x^{\alpha}$. This definition then implies that a power law is not a single function but an asymptotical composite function. The slowly varying function $L(x)$ can be thought of as the deviation from a pure power law for finite x.

For $f(x) = L(x)x^{\alpha}$, taking logarithms of both sides and dividing by $\log(x)$ gives

$$\log f(x)/\log(x) = -\alpha + \log L(x)/\log(x) \tag{4.28}$$

Remembering that $L(x)$ is a slowly varying function, in the limit, the second term on the right vanishes to zero as $x \to \infty$, and thus we have:

$$\log f(x)/\log(x) = -\alpha,$$

or equivalently,

$$f(x) = x^{-\alpha}, \text{ for } x \to \infty.$$

This means that the empirical form of the function becomes:

$$f(x) \sim x^{-\alpha} \tag{4.29}$$

and in terms of probabilities, a the cumulative function $P(S > x) = kx^{-\alpha}$ corresponds to a probability density function: $f(x) = k\alpha x^{-(\alpha+1)}$.

42 Note that this limit is not the one possible, but remains a realistic device, e.g., in finance.

2. Main Properties

We list below only properties directly related to two theorems proposed in this annex.

a) The property that most interests us and that generally makes power laws special is that they describe scale free phenomena. A variable undergoes a scale transformation of the form $x \to Cx$. If x is transformed, we then obtain:

$$f(x) = kC^{\alpha}x^{\alpha} = C^{\alpha}f(x) \tag{4.30}$$

provided that given initial power law function is $f(x) = kx^{\alpha}$. Changing the scale of the independent variable thus preserves the functional form of the solution but with a change in its scale. This is an important property in our case. Scale-free behaviour strongly suggests that the same mechanism is at work across different sectors of the economy, the industrial structure of which remains constant over a relatively long period of time, measured with any time measurement (i.e., seconds, minutes, hours, days, years). A useful example that should be appealing for economists is price. We say that price is a homogenous function of degree zero with respect to income.

b) A power law is just a linear relationship between logarithms (Breisinger et al., 2009) of the form:

$$\log f(x) = -\alpha \log (x) + \log k. \tag{4.31}$$

c) Power law also has excellent aggregation properties[43]. The property of being distributed according to a power law is conserved under addition, multiplication, polynomial transformation, min, and max. The general rule is that when combining two power law variables, the fattest power law (i.e., the one with the smallest exponent) dominates. This property could be helpful for empiricist researchers using this form of function.

Let $X_1,\ldots,X_n$ be independent random variables, and k, a positive constant. Let α_x be also the power law exponent of variable X. Following Gabaix (2008), Jessen and Mikosch (2006), we have the so-called inheritance mechanism for power law:

$$\alpha_{X_1+X_2.....X_n} = \min(\alpha_{X_1}, \alpha_{X_2}, \ldots \alpha_{X_n}) \tag{4.32}$$

$$\alpha_{X_1*X_2.....*X_n} = \min(\alpha_{X_1}, \alpha_{X_2}, \ldots \alpha_{X_n}) \tag{4.33}$$

$$\alpha_{\max(X_1,X_2.....X_n)} = \min(\alpha_{X_1}, \alpha_{X_2}, \ldots \alpha_{X_n}) \tag{4.34}$$

43 The interested reader is recommended to see the works of Jessen & Mikosch (2006) or Gabaix (2008). As an example of relative facilities of proofs, if $P(X > x) = kx^{-\alpha_x}$, then $P(X^k > x) = kx^{-\alpha_x} = P(X > x^{\frac{1}{k}}) = kx^{-\frac{\alpha_x}{k}}$, so $\alpha_{X^k} = \dfrac{\alpha_X}{k}$.

$$\alpha_{\min(X_1,X_2.....X_n})=(\alpha_{X_1}+\alpha_{X_2},..+.\alpha_{X_n}) \tag{4.35}$$

$$\alpha_{kx}=\alpha_x \tag{4.36}$$

$$\alpha_{X^k}=\frac{\alpha_X}{k} \tag{4.37}$$

Thus, if we have two variables X and Y with different exponents, this property holds when Y is normal, lognormal, or exponential, in which case $\alpha_Y = \infty$. Hence, multiplying by normal variables, adding non-fat tail noise, or summing over independent and identically distributed (i.i.d.) variables preserves the exponent.

This is a reason for hope for empiricists, that power law exponents can be measured even if the data are noisy. Although noise affects statistics (moments), it will not affect the PL exponent. The problem of missing data may not affect information contained inside data, either.

3. Statistical Complexity of a SAM

A social accounting matrix represents an economic table aggregating information about complex interchanges within different sectors and/or institutions. These interchanges have been described in Figure 1 where a general scheme of income flow in the economy was described. In economics, the main purpose of human activity is increasing income, the principal source of well-being. Changes in that income are usually assessed through gross domestic product growth (GDPG). However, this GDPG is itself an aggregate accounting of income growth from different sectors and institutions. Since a SAM is built under the principle of double entry bookkeeping, income, and expense totals should balance. This aspect has been previously alluded to. We remain within the Walrasian economy[44], which rules out that expenses growth is absolutely co-integrated, in the context of Granger time series analysis, with incomes or wealth, over any time period.

As earlier suggested, the economic system described is defined by different interactive subsystems, each represented by respective actors and characterized by optimizing behaviour. Households, which tend to maximize a certain utility function, remain the owner of factors of production and are the final consumer of produced commodities; firms maximize profits by optimal renting of these factors from households for the production of goods and services. In this model, government has the passive role of collecting and disbursing taxes. Furthermore, the economy analysed

44 However, according to recent research, power law consistency with equilibrium theory has, so far, failed to address this. Nevertheless, such consistency is expected for both theories (Doyne, & Geanakoplos, 2009) is small and open, and entirely prone to world fluctuations, owing to, among other things, the country's status of 'price taker' from the international market.

Furthermore, as already noted, due to different and sometimes contradictory sources of collected statistical information, a SAM cannot be balanced. Such statistical data may display, as partially coming from statistical surveys, systematic and stochastic errors, thus missing some normal Gaussian properties.

A Proof of SAM Power Law Distribution Properties

For the next step, we provide propositions evidenced by the above properties of power law functions and by other recent works in econophysics (e.g., Stanley et al., 1998).

Proposition 1. Under general Walrasian conditions,[45] the present level of sectorial (or institutional) income or expense total is a linear function of cumulated past and present sectorial or institutional wealth (income) or expenses growth rates of the global economy.

Proposition 2. Income (profit) growth rate follows exponential law within sectors or institutions with similar activity scale while this distribution becomes a power law among firms with different activity scales.

The first proposition simply explains a cumulative character of wealth from additional net incomes over time, in this case, sectorial industry or an institution. The property (4.32) guarantees plausibility of this proposition provided the independence of growth rates.

The second proposition, follows results found in the case of the U.S. economy (Stanley et al., 2001), where firms with the same level of activity display an exponential distribution of income (profit) growth rate while that distribution becomes a power law when we confront firms with different levels of activity. Furthermore, in this last case, the above authors have noted seven different fractals within that distribution.

Assumption 1 (structural stability). We will adopt an economy where factors of production are mobile among sectors of production and different scales of sectors are not affected by structural differences in factor productivities. To make this assumption more realistic, this means that the level of productivity and of factors within different sectors of the economy are identical enough so that there are no observed factor movements towards a given sector over a long period $T \to \infty$.

Assumption 2 (convergence). A cumulated economic growth rate trend is positive. This means that if we assume $g_{.jt}$ to be any growth rate in the economic or institutional sector *j* for period t, we have:

$$\sum_{t=1}^{\infty} g_{jt} \to \infty \qquad (4.38)$$

45 Here, we have particularly the principles of market *clearance* and of *income balance* in mind.

In other terms, positive growth rates always mark an advantage on negative ones and economic progress in the long run is guaranteed. This assumption is a stylized version of human economic development through history, owing to their natural capacity to innovate.

Theorem 1*. For a given non-centralized economy, disaggregated subaccounts of a vector (matrix) additively defining micro-elements of an entire system account by row or by column (being a cumulative income growth over finite lengths, periods n) display by row (expenses) or column (incomes) a power law distribution.*

Proof[46]: Let us first provide the demonstration on the income generating side of an accounts table and thus consider a non-centralized economy (system) made of *M* sectors and institutions (micro-elements) j (j = 1..M) generating each income w_{ijt} where *i* (*i* = 1…M) means one of the M sectors receiving incomes from one of the j sectors during the period t_n. Aggregative sectorial income is

$$\sum_{j=1..M} w_{\bullet j} = W_{\bullet}.$$

Let us consider two free periods of time 0 and t. We have then $w_{.j0}$ and $w_{.jt}$, two successive incomes during two periods, and we defined wealth growth g.j as g.j = $w_{.jt}$ / $w_{.j0}$, meaning a relative growth of wealth at period t. Equivalently, we have the growth rate ln(g.). In probabilistic terms, we will assume there is some collection $g_{.j}$ of possible wealth growth in a fixed sector i by a finite number of sectors j that each can generate with associated probabilities $\{p_{.j}\}$. For a fixed sector i, we have

$$\sum_{j=1..M} p_{\bullet j} = P_{\bullet} = 1 \tag{4.39}$$

Let us now introduce maximum entropy formalism to the problem.

Let us consider a continuous case where we have wealth growth g.j and its density of probability to be found, f(g.j). We maximize the entropy (Carter, 2011).

$$Max H(f) = -\int_1^{\infty} f(g_{\bullet j}) \ln((g_{\bullet j})) dg_{\bullet j} \tag{4.40}$$

subject to

$$\frac{r^{q-1}-1}{q-1} \geq 1 - \frac{1}{r} \tag{4.41}$$

$$\int_1^{\infty} f(g_{\bullet j}) \ln((g_{\bullet j})) dg_{\bullet j} = k \ln(g_{\bullet j}); \tag{4.42}$$

46 It will suffice to demonstrate the case of, for example, wealth growth and to deduct the case of expenditure growth, thanks to Walrasian aspects of our economy or by referring to the co-integrating character of both variables on a longer interval of time.

where constant k is the average number of inter-sectorial transactions per time step. The next step consists of applying the calculus of variations to maximize over a class of functions. Thus, solving an external problem of the functional:

$$\int F\,[z, f(z), f'(z)]dz \tag{4.43}$$

and look then at solving:

$$\frac{\partial F}{\partial f(z)} - \frac{d}{dz}\left(\frac{\partial F}{\partial f^{'}(z)}\right) = 0$$

We define the Lagrange of the form:

$$L \equiv -\int_1^{\infty} f(g_{\bullet j})\ln((g_{\bullet j}))dg_{\bullet j} - \mu\left(\int_1^{\infty} f(g_{\bullet j})dg_{\bullet j} - 1\right) - \lambda\left(\int_1^{\infty} f(g_{\bullet j})\ln((g_{\bullet j}))dg_{\bullet j} - k\ln(g_{\bullet j})\right)$$

Finally, we get from conditions of first order:

$$f(g_{\bullet j}) = e^{-(\lambda_0 - \lambda\ln(g_{\bullet j}))} = g_{\bullet j}^{\;-\lambda} e^{-\lambda_0}$$

where $\lambda_0 = 1 + \mu$.

One can use the normalization (4.41) condition to solve for $e^{-\lambda_0}$,

$$e^{-\lambda_0} = \int_1^{\infty} g_{\bullet j}^{\;-\lambda} dg_{\bullet j} = \left[\frac{g_{\bullet j}^{\;1-\lambda}}{1-\lambda}\right]_1^{\infty} = \frac{1}{\lambda - 1}$$

after assuming $\lambda \succ 1$.

Rearranging the above terms, we get the density probability functional form:

$$f(g_{.j}) = (\lambda - 1)(g_{.j})^{-\lambda} \tag{4.44}$$

This is the sought *density probabilities* of wealth growth rate of economic sectors and institutions. It displays a power law distribution form.

Next, Theorem 1 above and properties (4.32– 4.36) guarantee that the cumulated by the past growth sectorial incomes should continue to display a power law distribution irrespective of which form of transition economy evolves from period t to the next period t + 1. In particular, we note that property (4.33) ensures that multiplicative transitory combinations of different growth rates continue to keep the power law property of economic sectorial movements unmodified. Assumptions 1 and 2 guarantee that cumulated income growth rates are an increasing function of time, guaranteeing increasing sectorial wealth over generations.

The demonstration is proven.

Theorem 2. *Economic growth rate movements of any open economy display power law function properties.*

Proof: Its proof results from the demonstration of Theorem 1 and the additive property (4.32) since the global economic growth rate is derived as a weighted linear combination of sectorial growth rates. This ends the demonstration.

Bibliography – Part IV

Amaral, L.A., Buldyrev, S.V., & Havlin, M.A. (1998). Power Law Scaling for a System of Interacting Units with Complex Internal Structure. *Phys. Rev. Lett.* 80:1385–1388.

Ashok, P. (1979). Forecasts of Input-Output Matrices Using the R.A.S. Method. *The Review of Economics and Statistics,* 61(3):477–481.

Bacharach, M. (1970). *Biproportional Matrices and Input-Output Change*. Cambridge: University Press, Cambridge.

Bottazzi, G., Cefis, E., Dosi, G., & Secchi, A. (2007). Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics,* 29:137–159.

Breisinger, C., Thomas, M., & Thurlow, J. (2009). *Social accounting matrices and multiplier analysis: An introduction with exercises*. Food Security in Practice technical guide 5. Washington, D.C.: International Food Policy Research Institute.

Brown, A., & Stone, R. (1962). *A computable model for economic growth.* Cambridge, UK: Cambridge Growth Project.

Brown, D.J., & Shannon, C. (1997). Uniqueness, Stability, and Comparative Statics in Rationalizable Walrasian Markets. Cowles Foundation Discussion Paper No. 1170. Retrieved (1998 version): http://cowles.yale.edu/sites/default/files/files/pub/d11/d1170.pdf

Bwanakare, S. (2013). *Méthodologie pour la balance d'une matrice de comptabilité sociale par l'approche économétrique de l'entropie: le cas du Gabon*. Paris: Editions JePublie.

Bwanakare, S. (2014). Balancing an Ill-behaved National Accounts Table Under Power Law Property Hypothesis. *Polish Statistical Review,* 3:9.

Bwanakare, S., Cierpiał-Wolan, M., & Mantaj, A. (2016). Predicting Gross Domestic Product Components through Tsallis Entropy Econometrics, *Acta Physica Polonica A,* tom 129, nr 5, p. 993.

Carter, T. (2011). An introduction to information theory and entropy. [Online]. Available: http://astarte.csustan.edu/~ tom/SFI-CSSS.

Champernowne, D.G. (1953). A Model of Income Distribution. *The Economic Journal*, 63:318–351.

Doyne, J., & Geanakoplos, J. (2009). The Virtues and Vices of Equilibrium and the Future of Financial Economics, Cowles Foundation Paper No. 1274, In *Wiley Periodicals,* 14(3):11–38.

Duncan, K.F., & Smith, E. (2008). Classical thermodynamics and economic general equilibrium theory, *Journal of Economic Dynamics and Control,* 32:7–65.

Duncan, K.F. (1999). *Statistical Equilibrium in Economics: Method, Interpretation and an Example*. New School University.

Foley, D.K. (1994). A statistical equilibrium theory of markets. *J. of Econ. Theory* 62(2):321–345.

Gabaix, X. (2008). Power Laws in Economics and Finance. Retrieved: http://www.nber.org/papers/w14299

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

Graham, P. (1985). Commodity Balances and National Accounts – A SAM Perspective, *Review of Income and Wealth,* 2:155–169.

Jaynes, E.T. (1957b). *Probability Theory: The Logic Of Science.* USA: Washington University.

Jessen, A.H., Mikosch, T. (2006). Regularly Varying Functions. *Publications de l'Institut de Mathematique,* 80(94):171–192.

Kerwat, J., Dewhurst, J., & Molana, H. (2009). Constructing a social accounting matrix for Libya. (Dundee Discussion Papers in Economics; No. 223). University of Dundee. Retrieved: http://discovery.dundee.ac.uk/portal/files/108141/DDPE_223.pdf

Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics* 22:79–86.

Liossatos, P.S. (2004). *Statistical Entropy in General Equilibrium Theory,* Florida, USA: Department of Economics, Florida International University.

Mansur, A., Whalley, J, 1984, Numerical specification of applied general equilibrium models: Estimation, Calibration, and Data, in Scarf, H.E., and Shoven, J.B. (Eds.), 1984, *Applied General Equilibrium Analysis*, Cambridge, UK: Cambridge University Press.

Mantegna, R.N., Stanley, H.E. (2007). I*ntroduction to Econophysics*. Cambridge, UK: Cambridge University Press.

Miller, A.C., Matthews, A., Donnellan, T., & O'Donoghue, C. (2005). Social Accounting Matrix (SAM) for Ireland. Retrieved: http://ideas.repec.org/p/iis/dispap/iiisdp365.html.

Murat, A. (2005). Turkish Financial Social Accounting Matrix. Retrieved: http://ssrn.com/abstract=782729.

Pan, X., (2000). Social and Ecological Accounting Matrix: an Empirical Study for China, Beijing 100080: Chinese Academy of Sciences.

Plich, M. (2002). *Budowa i zastosowanie wielosektorowych modeli ekonomiczno-ekologicznych.* Łódź: Wydawnictwo Uniwersytetu Łódzkiego, 2002.

Pukelsheim, F. (1994). The Three Sigma Rule, Journal of the American Statistical Association, 48(2):88–91.

Pyatt, G. (1985). Commodity Balances and National Accounts – A SAM Perspective. In *Review of Income and Wealth,* vol. Series 31, no. 2, pp. 155–169.

Robinson, S., Cattaneo, A. & El-Said, M. (2001). Updating and Estimating a Social Accounting Matrix Using Cross Entropy Methods, *Economic Systems Research,* 1:47–64.

Salem, H.H., *The Macroeconomic Social Accounting Matrix of Tunisia in 1996*. G.A.I.N.S., Faculté de Droit et des Scieces Economiques, Université du Maine, Le Mans, France. Retrieved: https://www.gtap.agecon.purdue.edu/resources/download/1977.pdf

Scandizzoa, P.L. & C. Ferrareseb, C. (2015). Social accounting matrix: A new estimation methodology, *Journal of Policy Modeling,* 1:14–34.

Scrieciu, S.S., & Blake, A. (2005). General Equilibrium Modelling Applied to Romania (GEMAR): Focusing on the Agricultural and Food Sectors, Impact Assessment Research Centre, Working Paper Series, 11/2005, IARC, University of Manchester.

Stanley, E.H. (1998). Power Law Scaling for a System of Interacting Units with Complex Internal Structure. *Physical Review Leters,* 80(7):1385–1388.

Stone, R. (1970). *Mathematical Models of the Economy and Other Essays*. London : Chapman & Hall.

Toh, M.H. (1998). The RAS approach in updating input-output matrices: An instrumental variable interpretation and analysis of structural change, *Economic Systems Research,* 10(1):63–78.

Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. Springer, Berlin.

Wing, I.S. (2004). Computable General Equilibrium Models and Their Use in Economy-Wide Policy Analysis, Massachusetts Institute of Technology: Technical Note No. 6, MIT Joint Program on the Science and Policy of Global Change.

Zellner, A. (1991). Bayesian Methods and Entropy in Economics and Econometrics. In *The Netherlands*. (Eds. W.T. Grandy and L.H. Schick), pp. 17–31. Kluwer Academic Publishers, The Netherlands.

PART V: **Computable General Equilibrium Models**

1 A Historical Perspective

Since the 19th century, there has been a methodological discussion in economics about how one should analyse national economies. Walras (1834) introduced a framework he called General Equilibrium Analysis. According to Walras, in a national economy "everything affects everything." The only correct way to analyse it should be to treat the national economy as an unbroken entity and to use the tools of general equilibrium analysis. Later, many other economists have strengthened the Walras orientation. As an example, one can cite the Edgeworth diagram, which enabled us to explain the Pareto optimum. However, the work of Arrow and Debreu (1954) has constituted the cornerstone of Walrasian equilibrium theory. In fact, by proving the existence of the uniqueness of the optimum of Walrasian equilibrium—plus the two theorems of welfare—Arrow and Debreu combined abstract general equilibrium structure with realistic economic data to solve numerically for the levels of supply, demand, and price that support equilibrium across a specified set of markets. This allowed Walrasian equilibrium to become an applicable theory. It is worthwhile to recall here the contribution of Nash (1950), who introduced anticipation aspects into multi-game equilibrium, thereby achieving something like a quasi-Pareto optimum.

This school of thought is the forefather of the Leontief input-output model of production, the social accounting matrix (SAM), and microeconomic-based computable general equilibrium models (CGE).

On the other hand, Marshall (1890) criticized Walras and postulated that general equilibrium analysis is impossible in practice, because it demands too much information. Marshall's claim was that it is enough to separate from the rest of national economy the part under investigation and to analyse it within the framework he called *partial equilibrium analysis*. As a motivation, Marshall developed the so called *ceteris paribus* condition, which means "all other things remaining unchanged." Most post-Keynesian macroeconomic models belong to this school of thought. Criticism against large-scale macro econometric models built in the tradition of the Cowles Commission approach began in the late 1960s. These misgivings were subsequently reflected in the Lucas critique (parameters of models may take into account the reaction effect of agents with respect to expectation—rational or not), Sims's critique (time series models), and disenchantment with the model's Keynesian foundations (IS-LM models and the Philips curve) criticised by the Chicago school.

In response, classical macro econometric modelling progressed in two parallel ways: one, the improvement of the structure of traditional models, particularly in terms of specifying the supply-side and forward-looking expectations; and the other, strengthening techniques or developing alternative techniques (the so-called no economics theory-oriented models), e.g., the LSE approach aided by the advent of co-integration analysis, vector autoregressive (VAR) systems, and dynamic stochastic general equilibrium (DSGE) models.

Walrasian general equilibrium theory made its resurgence while the Keynesian model started declining. A major stimulus to early CGE modelling was Stone and Brown (1962). As a continuation of Leontiew's work (1941) — and to a certain degree, F. Quesnay's *tableau économique* (18th century) — Stone pioneered the development of the SAM framework with his 1955 article *Input-Output and Social Accounts* (1962). The general shape of a SAM framework was next described by Pyatt and Thorbecke (1976). Then, Pyatt and Roe (1977) published a book giving a detailed description of the example of Sri Lanka. Since then, SAMs have been applied in a wide variety of (developed and developing) countries and regions, and with a wide variety of goals, in particular, as we will see latter, for impact analysis and simulations.

While in the early 1960s CGE models were perceived as precious devises for modelling poorer economies (e.g., Adelman et al. (1978), Arrow et al. (1971), de Melo (1988)), CGE modelling of developed economies stems from Leif Johansen's 1960 sectorial growth model (MSG) of Norway as an extension of the Leontief model. The model was later extended by Harberger (1959, 1962). Showen, Scarf, and Walley (1984, 1972, 1992) with the presentation by Scarf (1969) of an algorithm helping to solve the model. Similarly, as far as CGE models for developed countries are concerned, since the early of 1960's, a model was developed by the Cambridge Growth Project under the initiative of Richard Stone in the UK. The Australian MONASH model is the next new generation representative of this class. Both models were dynamic (traced variables through time). Other more recent contributions may draw attention, in particular those of Jorgenson, using an econometric approach, Mc Kenzie (1959, 1981, 1987), Ginsburgh and Waelbroeck (1981, 1976), Ginsburgh and Keyzer (1997), Harris and Cox (1983), Bourguignon (1983), Decaluwe and Martens (1987, 1988). Today there are many other CGE models from different countries. One of the most well-known CGE models is the GTAP (Global Trade Analysis Project) model of world trade, which involves many researchers around the world.

Depending on, among other things, targeted time-scope analysis, the macroeconomics school of thought involved, or the approach to model estimation, nowadays there are large classes of CGE models. Readers interested in the epistemological aspects of CGE models can see—e.g., Xian (1984) Jorgenson (1984, 1998a), Ginsburgh and Keyzer (1997), McKenzie (1954) or Mansur and Whalley (1984). However, for the clarity of the document, let us now concentrate on two classes of CGE models. The first class models the reactions of the economy over a given perspective of time thus suggesting comparative static and dynamic CGE models. The second focuses on the theoretical aspects of equilibrium, seeing that economic conditions of general equilibrium are not always fulfilled.

As far as the first class of models is concerned, many CGE models around the world are static; that is, they model the reactions of the economy at only one point in time. For policy analysis, a simulation analysis is carried out and outputs are often interpreted as showing the reaction of the economy, in some future period, to one or more external shocks or policy changes. From the analytical point of view, the results show the difference (usually reported in percent of change) between two conditional alternative future states, that is, "what would happen if the policy shock were implemented." As opposed to dynamic models, the process of adjustment to the new equilibrium is not explicitly represented in such a model. However, details of the closure rule lead modellers to distinguish between short-run and long-run equilibriums. For example, this will be the case if the hypothesis on whether capital stocks are allowed to adjust or not.

Dynamic CGE models, by contrast, explicitly trace each variable through time at regular time steps, generally at annual intervals. While this class of model removes one of the main criticisms of CGE models, that of being unrealistic, as their analysis is based on one-year observations, at the same time, they become more challenging to construct and solve—they require, for instance, that future changes are predicted for all exogenous variables, not just those affected by a possible policy change. Furthermore, dynamic elements may arise from partial adjustment processes or from stock/flow accumulation relations—between capital stocks and investment and between foreign debt and trade deficits.

Recursive-dynamic CGE models are those that can be solved sequentially, over time. They assume that behaviour depends only on current and past states of the economy. The construction of this class of models is less complex and such models are easier to implement in empirical research than dynamic models.

Alternatively, if agents' expectations depend on the future state of the economy, it becomes necessary to solve for all periods simultaneously, leading to full multi-period dynamic CGE models. Recent publications cover this group of models, known as **dynamic stochastic general equilibrium** (DSGE) as they explicitly incorporate uncertainty about the future. It is worthwhile to add that the earliest DSGE models were formulated in an attempt to provide an internally consistent framework to investigate real business cycle (RBC) theory[47].

If we consider the second class of models focusing upon the general equilibrium aspects, one may consider that most CGE models rarely conform to the theoretical general equilibrium model. For instance, the presence of imperfect competition, non-clearing markets, or externalities (e.g., pollution) will lead the economy to disequilibrium conditions.

47 See DSGE: Modern Macroeconomics and Regional Economic Modeling by Dan S. Rickman, Oklahoma State University, prepared for presentation in the JRS 50th Anniversary Symposium at the Federal Reserve Bank of New York.

2 The CGE Model Among Other Models

We start this section with the next analytical question: How will a tax increase on petrol/gasoline impact the Polish economy? A tax of this kind would probably affect petrol/gasoline prices and might affect transport costs, car cost, the CPI, and hence, wages and employment. Traditional econometric models would have difficulty answering this question seeing the complexity of the tax shock response channels would imply on the whole economic system. CGE models are useful whenever we wish to estimate the effect of changes in one part of the economy upon the rest of the economy. They are being used widely to analyse trade policy. More recently, the CGE has been a popular device to estimate the economic effects of measures to reduce greenhouse gas emissions. Many research centres and central banks (including the Polish Central Bank) build stochastic dynamic CGE models which encompass the financial sphere of the economy.

A CGE model consists of behavioural equations describing model variables consistent with a relatively detailed database of economic information. The standard CGE tends to be neoclassical in spirit by assuming cost-minimizing behaviour by producers, average-cost pricing, and household demands based on optimizing behaviour. Thus, CGE models are built based on the theoretical model of competitive general equilibrium. Its original structure was developed during the second half of the 19th century by neoclassical economists. Among them, four should be mentioned here, the German Gossen (1983), the British Stanley Jevons (1879), the Austrian Menger (1871), and the French Walras. Due to the dominance of his contributions to its conceptualization, the model bears the name of the last author, the *Walrasian general system*.

3 Optimal Behaviour and the General Equilibrium Model

3.1 Introduction

The fact that the model is "computable" means that a numerical solution exists (e.g., Arrow-Debreu, 1954; McKenzie, 1959; Ginsburgh and Keyzer, 1997), and "general equilibrium" refers to simultaneously matching demand and supply on all markets.

In the example below, note the difference between a partial and a general equilibrium in the traditional way of analysing a market handed down by the Marshall and Walras schools. Let us suppose a Cobb-Douglas two-sector economy with two commodities X_i two sector inputs L_1,K_1 (labour and capital sectors) and two sector income Y_i, with i = 1,2. Then, the partial equilibrium model is defined by the next optimal program:

Objective:

$$Max_{x_1 x_2} \longrightarrow U = X^{\gamma}{}_1 X^{1-\gamma}{}_2$$

Market clearance:

$$Y_i = X_i \qquad\qquad i = 1{,}2$$

Production:

$$Y_i = A_i L_i^{\alpha_i} K_i^{1-\alpha_i}$$

Resource constraints:

$$L_1 + L_2 = \bar{L}$$
$$K_1 + K_2 = \bar{K}$$

In the case of a general equilibrium, we need to add an income balance restriction to ensure that all inflows and outflows are balanced.

Income balance:

$$P_1 X_1 + P_2 X_2 \leq w\bar{L} + r\bar{K}$$

with p_i (i = 1,2), w, r representing the prices of the two commodities, the sectors labour and capital respectively.

Let us now generalize the above formulation and consider a simple economy with m finite number of producers, n finite number of consumers, r commodities, and let us suppose that the Walras hypotheses are fulfilled. Thus, under these conditions, let us present, below, the behavioural functions of economic representative agents and conditions of market equilibrium.

Producer behaviour. Each producer, j (j = 1..m), is confronted with a set of possibilities of production v_j, the general element v_j of which is a program of production with dimension r, where outputs have a positive sign and inputs a negative sign. The objective of each producer is to select, for a given price p (p = 1..r), an optimal program of profits pv_j.

Consumer behaviour. Each consumer i (i = 1..n) is supposed to have an initial endowment of goods w_i (outputs or inputs) that the consumer is ready to exchange against remuneration by the producer.

Thus, the consumer is confronted with X_i possibilities of consumption of which the general element is x_i with dimension r. The consumer is never saturated in consuming X_i and his endowment w_i allows him to survive. For a given price p of dimension r, consumer i has the objective of maximizing total utility $U_i(x_i)$ under his given budgetary constraints:

$$pw_i + \sum_i \theta_{ij} pv_j = px_i \text{ with } x_i \in X_i$$

where θ_{ij} (i = 1,2 ..., n; j = 1,2, ..m) is a fraction of profits realized by the producer j and transferred to consumer i.

Producer function. *m* producers maximize individual total profits:

$$\text{Max} \quad pv_j = \tilde{p}\tilde{v}_j \tag{5.1}$$

subject to:

$$v_j \in V_j$$

Consumer function. n consumers maximize individual total utility:

$$\text{Max} \; U_i(x_i) = U_i(\tilde{x}_i) \tag{5.2}$$

subject to:

$$\tilde{p}x_i = \tilde{p}w_i + \sum \theta_{ij} \tilde{p}\tilde{v}_j \tag{5.3}$$

$$x_i \in X_i$$

This is definitely a general equilibrium solution $(\tilde{p}, \tilde{v}_j, \tilde{x}_i)$ from a decentralized system.

Market clearance. Excess demand for r goods is not positive:

$$\sum_i \tilde{x}_i - \sum_j \tilde{v}_j - \sum_i w_i \leq 0 \tag{5.4}$$

Commodities with supply excess, i.e., free commodities, have price zero while other commodities have a positive price:

$$\tilde{p}(\sum_i \tilde{x}_i - \sum_j \tilde{v}_j - \sum_i w_i) = 0 \tag{5.5}$$

where

$$pv_j = \tilde{p}\tilde{v}_j$$

A general equilibrium solution. This is definitely a general equilibrium solution $(\tilde{p}, \tilde{v}_j, \tilde{x}_i)$ guaranteeing that each of the markets will have realizable equilibrium. This, too, is an equilibrium for a decentralized economy since it guarantees compatibility of consumer and producer behaviours (Equations 5.1 and 5.2). This is a competitive equilibrium. The price from Equation (5.3) is imposed on all actors of the market.

3.2 Economic Efficiency Prerequisites for a Pareto Optimum

The purpose of this section is to clarify the connection between the general equilibrium model and the optimum Pareto state. This will allow us in the next chapter to go beyond such an equilibrium and to analyse impact on social welfare. We must then check whether or not the three conditions below are fulfilled.

a) ***Equality of marginal rates of technical substitution for different producers.*** Let us limit our generalization to an economy with two goods q_1 and q_2, and two limited inputs x_1 and x_1, for two respective producers.

$q_1 = f_1(x_{11}, x_{12})$, for producer 1,

$q_2 = f_2(x_{21}, x_{22})$, for producer 2,

This means that $\bar{x}_1 = x_{11} + x_{21}$ and $x_1 = x_{12} + x_{22}$
Let us maximize the quantity produced of q_1 under restriction of known quantity $\bar{q}_2$.
Using the Lagrange multiplier, we have:

$L = f_1(x_{11}, x_{12}) + \lambda[f_2(\bar{x}_1 - x_{11,} \bar{x}_2 - x_{12}) - \bar{q}_2]$

Finally, we get:

$$\frac{\frac{\partial f_1}{\partial x_{11}}}{\frac{\partial f_1}{\partial x_{12}}} = \frac{\frac{\partial f_2}{\partial x_{21}}}{\frac{\partial f_2}{\partial x_{22}}} = TmST_1 = TmST_2$$

The Pareto criterion having been satisfied, it becomes impossible to increase q_1 without decreasing q_2 and vice versa.

b) ***Marginal rate of substitution of products for different consumers.*** Let $U = f(q_1, q_2)$ be the total utility of any consumer and let q_1 and q_2 be the quantities consumed of two products. Assuming a constant level of total utility, the next relations follow:

$$dU = \frac{\partial U}{\partial q_1} dq_1 + \frac{\partial U}{\partial q_2} dq_2$$

$$-\frac{dq_2}{dq_1} = TmSP = \frac{\dfrac{\partial U}{\partial q_1}}{\dfrac{\partial U}{\partial q_2}}$$

where

$$\frac{\partial U}{\partial q_1}$$

and

$$\frac{\partial U}{\partial q_2}$$

are marginal utilities of the two goods.

As for the first condition, limiting our generalization to two consumers and two products which supply them, then one can pose:

$$U_1 = f_1(q_{11}, q_{12})$$

and

$$U_2 = f_2(q_{21}, q_{22})$$

U_1 and U_2 represent levels of utilities for the two consumers. The quantities q_1, q_2 are, respectively, consumed by consumer one and two. Thus, maximizing the utility of consumer 1 under the restriction of a given quantity of consumer 2 and using the Lagrange multiplier, we obtain:

$$L = f_1(q_{11}, q_{12}) + \lambda[f_2(\bar{q}_1 - q_{11,} \bar{q}_2 - q_{12}) - \bar{U}_2$$

and finally:

$$\frac{\dfrac{\partial f_1}{\partial q_{11}}}{\dfrac{\partial f_1}{\partial q_{12}}} = \frac{\dfrac{\partial f_2}{\partial q_{21}}}{\dfrac{\partial f_2}{\partial q_{22}}} = TmSP_1 = TmSP_2$$

Thus, the Pareto criterion is satisfied: it is impossible to increase U_1 without decreasing U_2 and vice versa.

c) ***About the marginal rate of transformation.*** The marginal rate of transformation of products is a measure in a global economy (and in absolute value) of how much supply of one product will increase as a consequence of an infinitesimal decrease in the supply of a second product.

We have:

$$-\frac{dq_2}{dq_1} = \frac{\dfrac{\partial q_2}{\partial x_{21}}}{\dfrac{\partial q_1}{\partial x_{11}}} = TmTP$$

Here the numerator and denominator of the second equality explain marginal physical productivities of inputs. To summarize conditions of attainment of the Pareto optimum or economic efficiency, we must have simultaneously fulfilled the three following prerequisites:

1. $TmST_1 = TmST_2 = \frac{r_1}{r_2}$ (r_i is the price of the input i)
2. $TmSP_1 = TmSP_2 = \frac{p_1}{p_2}$ (p_i is the price of the product i)
3. $TmTP = \frac{cm_1}{cm_2} = \frac{p_1}{p_2}$ (cm_i is the marginal cost of the product i)

   In the competitive market case, these three conditions (1, 2, 3 above) are simultaneously fulfilled, and we have:

4. $TmTP = TmST_1 = TmST_2 = \frac{p_1}{p_2}$

At the same time, this is a socially optimum Pareto. Resource allocation is optimal and leads to equality between marginal rate of substitution of products TmSP of consumer and the marginal rate of transformation between products TmTP inside the economy. Out of this optimal point, *better* for an individual would mean *worse* for another. In this context, a competitive market not only guarantees economic efficiencies but also social equity. We shall come back later to this aspect when we present some particular assumptions, shifting the economy from competitive market conditions towards a disequilibrium.

4 From a SAM to a CGE Model: a Cobb-Douglas Economy

4.1 CGE Building Steps

The CGE we want to present below belongs to the class of models described by Dervis et al. (1982) and remains similar to those reported by Robinson et al. (1990) and Devarajan et al. (1983). Thus the model is SAM based, which means that the SAM serves to identify the agents in the economy and provides a database with which the model is calibrated. As such, the modelling approach follows the influential Pyatt's *SAM Approach to Modelling* (Pyatt, 1988). In particular, since the model contains the important assumption of the law of one price, prices are common across the rows of the SAM. The SAM also provides an important organizational role since the groups of agents identified by the SAM structure are also used to define sub-matrices of the SAM for which behavioural relationships need to be defined. The model concerns a real economy without monetary sector.

As displayed by Figure 2, generating a CGE model proceeds in twelve steps (e.g., Decaluwé et al. 2001). However, since some of these steps are concerned mainly with data collection and processing, we describe below the six key steps directly related to model construction.

The first stage is the identification of the behavioural relationships. Since in the previous steps collected statistical data have been checked for their internal consistency with respect to macroeconomic theory[48], these are defined by reference to the sub-matrices of the SAM within which the associated transactions are recorded.

The second stage is formal and involves a further definition of the components of the transactions recorded in the SAM between respective sellers (rows) and buyers (columns). Thus, this step gives mathematical substance to behavioural relationships.

In the third stage, an algebraic statement of the general equilibrium model is provided. The equations of the model are succinctly presented in Figure 2, where standard accounts of the model's equations and variables are presented, too.

In the fourth stage, there is a discussion on macroeconomic closure rules and the choice of the most appropriate ones in the case of a modelled economy. This stage recognizes the fact that in CGE models, the number of variables is always higher than the number of equations, and exceeding variables must be rendered exogenous via appropriate macroeconomic theory.

48 However, as will be demonstrated by the extended input-output table or the social accounting matrix for taking into account environmental aspects in the economy, a SAM displays new internal consistencies, associating economy with environment.

1. Issue of the problem
2. Collecting statistical information
3. Checking internal consistency of collected data vs economic theory, e.g., a SAM must balance
4. Definition of behavioural relationships
5. Definition of transaction relationships
6. Algebraic statement of the model
7. Model closure rules
8. Choice of numerical values for the parameters of the model, i.e., calibration
9. Reproduction of the referential situation
10. Preparation of the simulation plan
11. Simulation and outputs of the new situation
12. Interpretation of outputs or comparison between referential and new situation results

Figure 2: The twelve steps for building a standard computable general equilibrium model[51]

The fifth stage consists of determining the numerical parameters of the model through the calibration technique. Then, some parameters not directly available must be derived from equations from the second stage so that the whole system of equations becomes determined. At this point, the application of the non-extensive entropy econometrics approach will reveal advantages over the traditional calibration technique. As shown in the preceding chapters, that new approach is conceptually better to solve inverse problems, nonlinear (e.g., the CES models) or not (e.g., the Cobb-Douglas models).

The last stage consists of replication (reproduction) of the initial situation through the already built model. When the model exactly replicates initial variable values contained within the SAM[49], the modeller can then provide various simulations to assess the macroeconomic impact of different shocks.

Behavioural relationships. The SAM accounts reflect the agents' behaviour related to different transactions that took place during a given period of time. A behavioural model is defined to describe the agent achieving each transaction posted in the SAM accounts. These behavioural relationships are defined by a set of linear and nonlinear functions reflecting the way agents involved in the model will respond to exogenous changes in the parameters and/or variables of the model. Next, the 'law of one price' should apply for each row of the SAM (McDonald, 2007), but this is not true in the present case.

49 This means that we have retained here the hypothesis that the SAM constitutes a basis for processing data.

50 This figure is an adaptation of Decalue et al. (2001) and Scot Mc Donald et al. (2007).

Households. Households maximize utility by choosing the bundles of consumption commodities through the Cobb-Douglas utility function type. Consumption commodities represent a set of composite goods and services from domestically produced and imported commodities. To reflect the hypothesis that home-produced and imported goods and services are imperfect substitutes, the above composite commodities are formally expressed as a Constant Elasticity of Substitution (CES) function. The optimal ratios of imported and domestic goods and services are set up through their relative prices. On the basis of the existence of product differentiation, Armington (1969) proposed the CES relation under the assumption of imperfect substitution (see Devarajan et al. (1983). According to this assumption, the same generic multiregional commodity is not perceived in the same way by different consumers. For instance, for a Polish consumer, Arabica coffee from Kenya may present a different taste from Arabica coffee from Brazil for various reasons. One assumes in this model the country to be a price taker for all imported goods and services.

Domestic production. A two-stage nested technology is used in the domestic production process. At the first level aggregate, one combines intermediate inputs with aggregate primary inputs, then generates the outputs of activities at the basic prices. Depending on the production technology chosen, the proportion of aggregate intermediates and aggregate primary inputs varies with the (composite) prices in the case of a CES specification while in the case of a Leontief model aggregate, intermediates and primary inputs are in fixed proportions, then reflecting a homogenous relation of degree one. The second level aggregate uses Leontief technology to generate intermediate input demands in fixed proportions with respect to aggregate intermediate inputs of each sector activity. The intermediate input commodity is a composite product which mixes domestic and imported raw materials. Also at next second level, primary inputs combine to generate aggregate value added through the Cobb-Douglas or CES technology. Once again, the relative factor prices determine optimal ratios of primary inputs. Each activity produces one aggregated commodity. Hence, each vector of goods and services demanded corresponds to one vector of activity outputs.

The vector of home commodities demanded is determined by the home demand (including self-consumption by households) and export, in both cases for home-produced commodities. Assuming imperfect transformation between home demand and export demand, formally explained by a Constant Elasticity of Transformation (CET) function, the relative prices on the alternative markets determine the optimal distribution of home-produced goods and services between home and export markets. In accordance with the country's position on the world export market, the model can be specified for a small country—and thus a price taker—or for a country with a dominant position—and thus a price-maker country—on all export markets. In fact, for a price-giver country, selected export commodities can be deemed to influence the world price.

The remaining behavioural relationships constructed on the basis of SAM accounts remain generally linear. Those concern principally inter-institution exchanges. Thus, we have tax rates, i.e., transfers from households or enterprises to government; social allocations, i.e., secondary distribution of income; and reciprocal transfers between enterprises and households or between local institutions and the rest of the world. To this list, which is not exhaustive, one can also add fixed rate of savings or investment.

4.2 The Standard Model of Transaction Relationships

In the standard model description below, we follow—for clearer presentation—symbol indexes found in McDonald and Thierfelder (2007) unless the context requires some changes. The transaction relationships are presented in Fig 3. The prices of domestically consumed (composite) commodities are defined as *PQDc*. This price is weighted prices of domestic and imported price at *CIF* (cost insurances and freight) value to which are added indirect taxes, e.g., value added tax (*VAT*). The quantities of commodities demanded domestically are, as expected, divided between intermediate demand, *QINTDc*, and final demand, with final demand further subdivided between transactions by households, *QCDc*, enterprises, *QENTDc*, government, *QGDc*, investment, *QINVDc*, and stock changes, *dstocconstc*. When no analytical need is explained, these last two categories of final demand are grouped together under investment. The value of total domestic demand, at purchaser prices, is therefore *PQDc* **QQc*. As already explained, export demand, *QEc*, as an entry in the commodity row is not taken into account here since the domestic prices of exported commodities, *PEc* = *PWEc* * *ER*, suggest a separate row for export commodity according to the law of *one row one price*. Here *PWEc* means world prices and *ER* exchange rate. Abbreviations are merely a way to save space in the table. Price of exported commodities is additionally affected by export duties, *TEc*, which are entered into the commodity columns. Commodity supplies come from domestic producers who receive the prices, *PXCc*, for each commodity *c*. with the total sectorial domestic production of commodities being denoted *QXCc*. Commodity imports, QMc, are valued carriage insurance and freight (*CIF*) paid, in a way that the domestic price of imports, PMc, is defined as a world price, PWMc, times the exchange rate, *ER*, plus an ad valorem adjustment for import *TMc*.

All domestically consumed commodities are subject to a variety of indirect taxes, such as sales taxes, *TSc*, and excise taxes, *TECc*. Other taxes can be readily added.

We assume that each activity produces a unique output *QXa*. Domestic production activities receive average prices for their output, *PXa*, being an aggregated price of intermediary input price and primary factors price. This means that in addition to intermediate inputs, activities also purchase primary inputs, *FDfa*, for which they pay average prices, *WFf*. The model allows the prices of each factor to differ among activities. Some or all activities pay production taxes, at the rates *TXa*, proportionately to the value of activity outputs.

	Enterprises	Government	Capital	RoW	Total
Commodities	$(PQD_c * QENTD_c)$	$(PQD_c * QGD_c)$	$(PQD_c * QINVD_c)$ $(PQD_c * dstocconst_c)$	$(PWE_c * QE_c * ER)$	$(PQD_c * QQ_c)$
Activities	0	0	0	0	$(PX_a * QX_a)$
Factors	0	0	0	$(factwor_f * ER)$	YF_f
Households	$hoentconst_h$	$(hogovconst_h * HGADJ)$	0	$(howor_h * ER)$	YH_h
Enterprises	0	$(entgovconst * EGADJ)$	0	$(entwor * ER)$	$EENT$
Government	$(TYE * YE)$	0	0	$(govwor * ER)$	EG
Capital	$(YE - EENT)$	$(YG - EG)$	0	$(CAPWOR * ER)$	$TOTSAV$
Rest of World	0	0	0	0	Total 'Expenditure' Abroad
Total	YE	YG	$INVEST$	Total 'Income' from Abroad	

Figure 3:The standard model transaction relationships.

Source: PROVIDE project, 2009 *Enterprises*

The model allows for the domestic use of both domestic and foreign owned factors of production and for payments by foreign activities for the use of domestically owned factors. Factor incomes, therefore, accrue from payments by domestic activities and foreign activities, *factworf,* where payments by foreign activities are assumed exogenously determined and are denominated in foreign currencies.

After allowing for depreciation, *deprecf*, and the payment of factor taxes, *TFf,* the residual factor incomes, *YFDISTf*, are divided between domestic institutions (households, enterprises, and government) and the rest of the world in fixed proportions. Households receive incomes from factor rentals and/or sales, inter-household transfers, *hohoconsthh*, transfers from enterprises, *hoentconsth*, and government, *hogovconsth*, and remittances from the rest of the world, *howorh*, where remittances are defined in terms of the foreign currency.

Household expenditures consist of payments of direct/income taxes, *TYh*. The household income after tax deduction is then posted to inter-household transfers and consumption expenditures, with the pattern of consumption expenditures determined by the household utility functions. The residual household income is posted to savings, where the saving rate, *SHHh*, is an exogenously fixed coefficient.

The enterprise account receives income primarily from capital returns in the form of retained profits. Depending on the scale and field of activity, the enterprise may receive subsidies from government, *entgovconst*, or transfers from the rest of the world, *entwor*.

Expenditures then consist of the payment of direct/income taxes to government, *TYE*, transfers of dividends, and other remittances to households, *hoentconsth*. Enterprises may transfer to the rest of the world a fraction of their profits or pay for foreign factor. Residuals which are the difference between income *YE* and expenditures *EENT* constitute the source of savings for enterprises.

Government. Government income comes primarily from direct and indirect taxation of factors, commodities, and from various penalties related to public administration. Thus, incomes accrue from the various tax instruments (import and export duties; sales, production, and factor taxes; and direct taxes), that can all vary due to changes in the values of production, trade, and consumption. The government may receive transfers (in foreign currency) from the rest of the world, *govwor*, as a consequence of various international conventions.

Expenditures consist of transfers to enterprises for subsidies of some productive activities. Other transfers concern the secondary distribution of income by government and are oriented toward households. Transfers may concern the rest of the world, particularly in the form of salary packages to nationals residing out of the country. In these models, all transfers to different domestic institutions are fixed in real terms but may vary in nominal terms with consumer prices. There is an analogous treatment of government savings as in the case of the enterprise, i.e., the internal balance, which is defined as the difference (residual) between government income, *YG,* and committed government expenditure, *EG*. The quantities of commodities consumed by the government are fixed in real terms, and hence government consumption expenditure will vary with commodity prices. In some circumstances, it can be realistic to endogenize government consumption behaviour for the determination of government consumption expenditure.

Domestic investment. Domestic investment demand consists of fixed capital formation, *QINVDc*, and stock business changes, *dstocconstc*. The departure point is that in an open economy model (an economy with foreign trade and capital flows), private investment plus governmental borrowing must equal private savings plus foreign investment. The value of fixed capital formation will vary with commodity prices while the volume of fixed capital formation can vary both as a consequence of the volume of savings changing or changes in exogenously determined parameters. In this basic version of the model, domestic savings is made up of savings by households, enterprises, the government (internal balance), and foreign savings, i.e., the balance on the capital account or external balance, *CAPWOR*.

Savings and capital accounts. The various closure rules available within the model allow for alternative assumptions about the determination of domestic savings. One could make assumptions about, for instance, flexible vs. fixed savings rates for households or flexible vs. fixed exchange rates for the value of foreign savings. Foreign savings come out as a discrepancy between foreign incomes and expenditures. The incomes to the rest of the world account, i.e., expenditures by the domestic economy towards the rest of the world, consist of the values of imported commodities and factor services. On the other hand, expenditures by the rest of the world account, i.e., incomes to the domestic economy from the rest of the world, consist of the values of exported commodities and net transfers by institutional accounts. The difference between these incomes and expenditures represent the capital account, the value of which depends on a variable exchange rate.

Prices. As in any standard model, the supply prices of the composite commodities *PQSc* are defined as the weighted averages of the domestically produced and consumed commodities *PDc* and the domestic prices of imported commodities *PMc*. Domestic prices of imported commodities are defined as the products of world prices of commodities *PWMc* and the exchange rate *ER* increased by ad valorem import duties *TMc* and by a consumption tax applied to domestic commodities, such as sales taxes *TSc* and excise taxes *TEXc*. These weights are updated in the model through first order conditions for optima. The above average prices will give the composite consumer price *PQDc*.

The producer prices of commodities *PXCc* are defined as the weighted averages of the prices received for domestically produced and sold commodities *PQSc* and export *PEc* markets. These weights are updated in the model through first order conditions for optima. Given that the country is by hypothesis small, the prices received on the export market are defined as the world price of exports *PWEc* and the exchange rate *ER* less any export duties due, which are defined by ad valorem export duty rates *TEc*.

The average price per unit of output received by each activity *PXa* is defined as the weighted average of the domestic producer prices. After paying indirect/production/output taxes *TXa*, this is divided between payments to aggregate value added *PVAa*, i.e., the amount available to pay primary inputs and aggregate intermediate inputs *PINTa*. Total payments for intermediate inputs per unit of aggregate intermediate input are defined as the weighted sums of the prices of the inputs *PQDc*. Let us recall that input prices are explained in composite price since imported commodities are demanded as intermediate inputs.

General Equilibrium conditions. As already mentioned several times, total demand for the composite commodities, *QQc*, consists of demand for intermediate inputs, *QINTDc*, consumption by households, QCDc, enterprises, *QENTD*c, and government, *QGDc*, gross fixed capital formation, *QINVDc*, and stock changes, *dstocconstc*.

Total supply, that is, supplies from domestic producers, *QDc*, plus imports, *QMc*, meets these demands; equilibrium conditions ensure that total supply and demand for all composite commodities must balance. Commodities are delivered to both the domestic and export, *QEc*, markets subject to equilibrium conditions that require all domestic commodity production, *QXCc*, to be either domestically consumed or exported.

As far as equilibrium of factor market *WF* is concerned, supply must balance with demand. In the case of the hypothesis of an inter-sector non-mobility of factors, equilibrium will have a sector character. The next balance concerns the current-account balance for the rest of the world in foreign currency. It implies that import spending plus factors to the rest of the world must balance with export income, plus revenues, institutional transfers from the rest of the world, and foreign savings. Next, equilibrium concerns the government balance, which means that government income *YG* must balance with government expenditures *EG* plus government savings *GSAV*. Finally, the savings-investment relation must balance, which means that non-government savings, plus government savings and foreign savings, must balance with fixed investment plus stock change.

Standard Mathematical formulation. Standard mathematical formulation of a CGE model is presented in almost each presentation related to this class of model, and interested readers can refer to Lofgren et al. (2000) or Herault (2006) amongst others. However, in the present model, we add some changes concerning the estimation of behavioural parameters to react to one of the main criticisms of CGE models. Production relationships by activities are defined by a series of nested Constant Elasticity of Substitution (*CES*) production functions. In the standard model, there is, however, additional limits imposed by economic meaningfulness and the availability of empirical data that allow for the inclusion of information (elasticity of substitution) about the possibilities for substitution between and within sub-groups of factors. This point is one of the repeated criticisms of CGE models since, in this particular case, different modellers should find different outputs, as the selection of elasticity parameters is arbitrary. In the present work, we bypass this problem thanks to a new maximum entropy approach introduced in the first parts of this document. The same procedure will be carried out for the constant elasticity of transformation between domestic and export product or Armington elasticity of substitution between domestic product and imports. Coming back to activity output, it is worth reminding that it is a CES aggregate of the quantities of aggregate intermediate inputs *QINT* and value added *QVA* while aggregate intermediate inputs are a Leontief aggregate of the (individual) intermediate inputs. Furthermore, aggregate added value is a CES aggregate of the quantities of 'primary' inputs demanded by each activity, where the primary inputs can be natural factors—types of labour, capital, and land. In the traditional model, optimal combinations of each natural and/or aggregate in each *CES* aggregate are determined by first order conditions based on relative prices.

5 Estimating the CGE Model through the Maximum Entropy Principle

5.1 Introduction

In their reference work, Arndt et al. (2002) present a new approach to estimating parameters of a CGE model through maximum entropy. That approach is pursued here. CGE models are frequently criticised for resting on weak empirical foundations (e.g., Shoven and Whaley, 1992). Whatever class of CGE model is in use, it displays consistent drawbacks, such as the lack of efficient methodology for the estimates of behavioural parameters (e.g., trade parameters), less realistic economic assumptions (e.g., representative agent), less flexibility in implementing the monetary sector, imperfect competition, and the impossibility of inferring through interval confidence. In Social Accounting Matrix (SAM)-based CGE model estimation, problems related to calibration to a benchmark period, the often not updated information lying in the input-output matrix—the principal part of a SAM—and various other criticisms, appear in the economic literature. Nevertheless, as is often underscored there, the problem of estimates of behavioural parameters is common to competing time series econometric models unable to predict future agent behaviour (see the Lucas critique, 1976). Rational expectation-based models (Muth, 1961 and Kydland et al., 1977), like the dynamic stochastic general equilibrium models (DSGE) (Kydland et al., 1982), have tried to overcome the problem. Nevertheless, the DSGE models continue to display conceptual drawbacks from different sources, like the limited knowledge about the data generating system—and its future evolution—to which parameters are related (Evans and G. Ramey (2006), Tovar (2009)). Additionally, Sims (1987) reports additional drawbacks to the 'Lucas critique' in the context of rational expectations. Most of time, this may suggest that through the model with full information, we will deal with an over-parameterized, non-ergodic inverse problem, which traditional econometric approaches have failed to handle.

Next, in most developing countries, the alternative macro econometric models display serious weakness owing to scarce statistical data over large periods of time leading to technical problems of estimation (the degree of freedom) (Arndt, 2002).

Authors, such as Guerrien (2000) have expressed their scepticism about the usefulness of CGE models. Bernard Guerrien vigorously points to the less realistic economic assumptions such as the maximizing utility representative agent, perfect competition (Arrow-Debreu frame) or even imperfect monopoly markets; all these neo-classical concepts are described as fictive reality. The above criticisms have put into question the last strength of CGE models, that of being built on microeconomic foundations. Technical approaches to respond to most of the above criticisms have been attempted over the last two decades. After a thorough comparison between CGE and competing macro or microeconomic models (fixed-price models, dynamic optimisation/optimal control models, macro econometric models), authors such as Capros (Capros et al., 1990) suggest that all these models are complementary. More interestingly, the same authors provide useful insights into a future model that could address the cited shortcomings of current models.

Reconciling opposing models built upon contradictory assumptions would constitute too great a challenge. An acceptable solution should probably consist in setting up a model with only limited and realistic assumptions, basically combining common advantages of a CGE and a macro econometric equilibrium model.

This seems to be in line with two groups of authors in their recent works. One is by Robinson (Arndt et al., 2001, Go et al., 2015) who proposed a parameter estimation for a CGE using Shannon-Gibbs maximum entropy econometrics. These authors list additional qualities in comparison to the classical approach. One may cite the incorporation into the model of prior information related to present or past periods, thereby introducing dynamic elements into the system; the rationale for using entropy econometric formalism is the already cited quality of performing well in the absence of copious data. Last, the approach gives quantified information on the capacity of the model to reproduce a statistical record and computes statistical significance of parameter estimates.

The second work is by Francois (2001) who tries to overcome the problem of calibration: during the estimation process, the base period values are used to set up initial variable levels for the next steps of numerical estimation. However, numerical, successive approximations at the end can generate important deviations from the true values particularly when numerical processes imply multiplicative intermediary errors.

In this chapter, we try to extend the approach of Robinson (Arndt et al., 2002) to one of non-extensive entropy to estimate behavioural parameters for a CGE model.

The above authors have emphasized that the maximum entropy approach is similar to the econometric approach of Jorgenson (1984, 1998a) in different aspects. To a certain extent, the full historical record can be employed and statistical tests for estimated parameter values are available. Furthermore, as pointed out in different works (e.g., Golan et al., 1996), the ME approach can be applied in the absence of copious data. The ME approach allows one to use all available data, take into account all relevant constraints, employ prior information about parameter values, and apply variable weights to alternative historical targets. Available information does not need to be complete or even internally consistent. The philosophy of the ME approach is to use all available information while avoiding the use of information not available, for example, strong assumptions about the distribution of error terms.

5.2 Estimation Approach

In line with the preceding section, here the central question is why propose a maximum entropy econometrics related approach among many other econometrics techniques conceived for modelling large scale macroeconomic phenomena. On this point, it is worthwhile to refer, among a vast literature on the subject, to the review work of the Nobel laureate C. A. Sims (2007) on Bayesian methods applied in econometrics, in which he explains *Why Econometrics Should Always and Everywhere Be Bayesian* and then rejects frequentist asymptotics-based econometrics techniques. However, the above author did not allude to the maximum entropy econometrics approach as a competitive approach to Bayesian models[51]. According to Sims, the Bayesian approach uses the "Bayes rule" to incorporate the present beliefs (prior) about the phenomenon and to update this information with new facts at hand (data). Thus, following this author, unlike the frequentist-related econometrics methods, "Bayesian Inference is a way of thinking, not a basket of 'methods'" (Sims, 2007).[52] It becomes interesting now to compare the Bayesian approach with the maximum entropy approach. Both approaches use priors and data to produce model posteriors. Nevertheless, beyond this similitude, the approaches are conceptually different.

Basically, applying Bayesian theorem means (Jaynes, 1988) just computing a probability and not a probability distribution. It follows that the Bayes theorem does not make any reference to sample space or hypothesis space. In empirical application of the Bayesian approach, we need to go beyond the "exploratory phase" to the point where a certain structure (the likelihood) of the model can be assessed through additional model distribution hypotheses.

On the contrary, the maximum entropy[53] approach requires us to define a hypothesis space which sets down the possibilities to be further considered. Thus, in the exploratory phase of the problem, one can apply the entropy principle to solve a problem. In recent work, Giffin (2009) compared through illustrative computa-

51 Nevertheless, he pointed out that the Shannon mutual information approach has a more limited estimating capacity than the Bayesian one.

52 Jaynes [59] underscores the importance of this debate: "The recent literature has many attempts to clarify the relation of these principles." Williams (1980) sees Bayes's theorem as a special case of MAXENT while van Campenhout & Cover (1981) see MAXENT as a special case of Bayes's theorem. In our view, both are correct as far as they go, but they consider only special cases. Zellner (1987) generalizes Williams's results: "Thus Williams considers the case where we have a set of possibilities (H1 .. Hn), and some new information E confines us to a subset of them. Such primitive information can be digested by either Bayes' theorem or MAXENT, leading, of course, to the same result; but Bayes' theorem is designed to cover far more general situations. Likewise, van Campenhout & Cover consider only the Darwin-Fowler scenario; MAXENT is designed to cover more general situations, where it does not make sense to speak of 'trials'".

53 To be more precise, we have in mind the method of relative Entropy (or cross-entropy) of which maximum entropy can be seen as a particular case.

tions relative entropy, maximum entropy, and the Bayes rule in the environment of moments and data. Giffin underscored his finding related to the relationship between these three approaches. Only relative entropy can solve a problem by the simultaneous processing of data and moments—what Bayes and Maximum entropy alone cannot handle.

An interesting comparison between the Bayesian approach and the generalized maximum entropy approach (GME) is presented in Golan et al. (1996). As reported in these authors' work, Casella and Strawdermann (1981) used a simple example of recovering a bounded mean from a single observation $x = \beta + e$, where $e \sim$ `$N(0,1)$ and $\beta \subset [-c, c]$ are unknown to recover an image of β from x. They first specified a discrete distribution with equal mass on points $-c$ and c, using the normal likelihood function and provided the Bayesian posterior mean—under the squared error loss—of the form:

$$\hat{B}_{\beta} = c \tanh(cx)$$

The above authors in Golan et al. (1996) performed the estimation of the same model using the generalized maximum entropy estimator and found the following GME solution:

$$\hat{B}_{GME} = c \tanh(-c\lambda)$$

where λ is the optimal Lagrange multiplier on the model constraint.

Authors showed that both estimators are related in mathematical formulation. Their estimates are equal only if $\lambda = -x$, which occurs when $x = 0$. In this basic example, we notice that the Bayesian solution is directly influenced by the observation data x while the GME solution, by the optimal Lagrange multiplier λ. Thus, the two formula illustrate not only the mutual consistency of both estimators but also a rich source of their confusion in scientific literature. The fact that the GME solution is a function of the optimal λ is of high interest as this parameter is a function of the whole model, including the moments and the maximum entropy objective function. In the case of the generalized cross-entropy (GCE) technique, the advantage of that parameter for defining the optimal solution of a model is much more evident. In such a case, the optimal λ displays a direct relationship with the prior, the data, the stochastic random disturbance, and the maximum entropy intrinsic properties. This may explain why the GCE approach seems to be preferred for solving stochastic ill-behaved inverse (non-ergodic) problems. To be more precise, the more constraining data are consistent with the model, the greater the value of λ (absolute value), and the less uniform the maximum cross-entropy probability distribution. Unlike the Bayes approach, the GCE approach does not need any imaginative theoretical hypothesis to create a solution space closed-form. Golan (*et al.*) [61] comparing performances of different classes of econometric estimator in the case of ill-conditioned problems underscored the highest solution stability provided by the GCE technique.

In fact, unlike other estimators, the penalty for using wrong prior information is much smaller relative to competitive estimators, including the Bayesian estimator. According to the above explanations, such superiority results from a diversified relationship between the optimal λ and other constraints (data, stochastic random disturbance, and the maximum entropy objective function): the GCE estimates should not go far from the true parameters if only one constraint—in this case the prior—is violated. This result points to the GCE as the best estimation approach, at least in the case of stochastic ill-behaved inverse problems, as has been confirmed in a recent study (Bwanakare, 2014) which estimated parameters of three classes of constant elasticity of substitution models. The performance—measured on the parameter error variance coefficient—of the Tsallis related non-extensive cross-entropy estimator was much higher in comparison with the traditional econometric techniques (LS, ML, GMM, and NLS).

The number of authors who have tried to link CGE models with the cross-entropy econometrics approach is still limited (e.g., Arndt et al., 2002, Judge and Mittelhammer, 2012, and Go et al., 2015). In Arndt (2002) and Go et al. (2015), the authors start by viewing a classic, static CGE model in the following form:

$$F(X,Z,B,\delta) = 0 \tag{5.6}$$

where:
F: an I-dimensional vector valued function,
X: an I-dimensional vector of endogenous variables, usually *prices* and *quantities*,
Z: a vector of exogenous variables such as endowments and tariff rates,
B: a K-dimensional vector of behavioural parameters such as unknown constant elasticity of substitution parameters,
δ: a second vector of behavioural parameters whose values are uniquely implied by the choice of B, the exact form of F, and data for the base year.

The elements of F capture production and consumption behaviour which is coherent in terms of economics as well as macroeconomic constraints. After parameter calibration and estimation, static CGE analysis proceeds by changing the vector of exogenous variables, Z, and examining through simulation the resulting vector of endogenous variables, X, which satisfies Equation (5.6).

In the entropy estimation formulation proposed by the above authors, the static model attempts to track the historical record over T ($t = 1,2,\dots,T$) time periods. The Z vector is partitioned into exogenous variables observable from historical data, Z_t^0, and exogenous variables not observable from historical data, Z_t^u. The vector Z_t^0 would typically contain historical data on elements such as tax rates, endowments, world prices, and government spending. The vector Z_t^u might contain rates of technical change, implicit or unknown tax or subsidy rates. These variables and other items are not available from the historical record and must be estimated. Due to calibration to the base year and the restrictions imposed on the function, F, a unique relationship between δ and B exists which permits the model in Equation (5.6) to reproduce the base year, conditional on the choice of behavioural parameters B,

$$\delta = \Phi(Z_{t'}, B) \tag{5.7}$$

Note that the full vector Zt′ is assumed observable in the base year, labelled year t′. Estimation occurs in the context of the CGE model, and then we have the next relationship:

$$F(X_t, Z_t^0, Z_t^u, B, \delta) = 0, \ \forall t \in T \tag{5.8}$$

having to hold for estimated values B and Z_t^u, imposed values Z_t^0, and calibrated values δ. The outputs from such a solved problem lead to a predicted historical time path for variables of interest. These time series outputs can be compared with actual historic time paths in the following way:

$$Y_t = G(X_t, Z_t^0, Z_t^u, B, \delta) + e_t \tag{5.9}$$

where:
Y_t : is an N-dimensional vector of historical targets defined inside the social accounting matrices,
G: is a functional producing the vector of model predicted values for the targets,
e_t: is an N-dimensional vector representing the discrepancy between historical targets representing the unknown data generating system F and predicted values evaluated by the functional G, using sample information. Calibration to the base year implies that $e_t' = 0$.

After reparametrization of parameters B, e_t, Z_t^u, on defined support spaces according to the methodology explained earlier, the authors propose to set up a Gibbs-related cross-entropy (CE) model to be minimized under restrictions presented above. Since we specify prior distributions on parameters, the objective contains the two terms, precision and prediction (Golan et al., 1996), and each term can be given a weighting factor, α_1 and α_2. This CE formulation may be written as follows:

$$\underset{p,r}{Min} \ \alpha_1 \sum_{k=1}^{K} \sum_{m=1}^{M} p_{km} \log\left(\frac{p_{km}}{q_{km}}\right) + \alpha_2 \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{j=1}^{J} r_{tnj} \log\left(\frac{r_{tnj}}{s_{tnj}}\right) \tag{5.10}$$

Subject to:

$$F(X_t, Z_t^0, Z_t^u, B, \delta) = 0, \quad \forall t \in T$$

$$Y_t = G(X_t, Z_t^0, Z_t^u, B, \delta) + e_t \quad \forall t \in T$$

$$\delta = \Phi(Z_{t'}, B)$$

$$B_k = \sum_{m=1}^{M} p_{km} v_{km} \quad \forall k \in K$$

$$e_{tn} = \sum_{j=1}^{J} r_{tnj} w_{tnj} \quad \forall t \in T, n \in N$$

$$\sum_{m=1}^{M} p_{km} = 1 \quad \forall k \in K$$

$$\sum_{j=1}^{J} r_{tnj} = 1 \quad \forall t \in T, n \in N.$$

Once again, the CE formulation in Equation (2.22) corresponds to the Kullback-Liebler measure of deviation of the estimated weights from the prior. The constrained optimization problem in Equations (2.23–2.24) chooses distributions for parameters and error terms that are closest to the prior distributions, using an entropy metric, and satisfies the full set of conditions required by a CGE model. In addition, the model endogenously calibrates itself to the base year.

The cited authors (Arndt et al., 2002) provide a case study on Mozambique.

Now, to extend the above approach to non-ergodic systems, we replace the objective function in (2.22) by the previously introduced criterion function of the form[54]:

$$\begin{aligned} Min \xrightarrow[p,r]{} \alpha_1 \frac{1}{q-1} \sum_k \sum_m p_{km}\left[(p_{km})^{q-1} - (q_{km})^{q-1}\right] - \sum_k \sum_m (p_{km} - q_{km})(q_{km})^{q-1} + \\ \alpha_2 \frac{1}{q-1} \sum_t \sum_n \sum_j r_{tnj}\left[(r_{tnj})^{q-1} - (s_{tnj})^{q-1}\right] - \sum_t \sum_n \sum_j (r_{ttnj} - s_{tnj})(s_{tnj})^{q-1} \end{aligned} \tag{5.11}$$

subject to the same restrictions as above in (5.10).

In this case, we have then assumed a higher complexity of statistical data-generating system, and/or other kinds of systematic errors to which collected data might be prone. This case recalls characteristics of power law discussed in previous chapters. The value of the Tsallis parameter q will inform us about the complexity of the system, as we already know.

The CGE outputs presented at the end of this chapter only limit entropy application to functions related to estimating behavioural parameters of constant elasticity of substitution. Further research on the methodology presented by the above authors could be of high interest. In particular, testing the proposed non-extensive relative entropy above, under the hypothesis of power law characteristics of macroeconomics remains urgent. In fact, in a recent publication, Bwanakare [62] has shown that trade functions used in CGE models may belong to the class of power law (Levy's process)

---

54 Note that there are two forms of Tsallis-Kullback-Leibler relative entropy. The one presented here comes from Bregman. Thus, it does not require *escort distribution* in the constraining block.

distribution. The implication is that not taking this into account by the Shannon-based entropy could lead to significant estimation errors. To prove this, we are going to provide below an example showing that only estimation of trade parameters of a CGE model under power law hypothesis leads to adequate outputs.

5.3 Application: Non-Extensive Entropy and Constant Elasticity of Substitution-Based Models

The example below shows the possibility of carrying out robust estimation of a stochastic constant elasticity of substitution (CES)-based model through the Tsallis entropy econometrics technique. The estimator properties for Tsallis entropic form have been suggested in Part II of this book, in the section devoted to the parameter confidence interval area-based statistical inference.

The technique presented below has been suggested by Bwanakare (2014) or more recently in Bwanakare (2016). This extended the results through the case study of eight CES production (CESP) models of seven countries. While the proposed approach could be generalized to a large class of nonlinear models, the example focuses on CGE trade models, the parameters of which are usually obtained through a calibration technique. The CES-based models remain intractable while trying to analytically estimate their parameters. In empirical research, various nonlinear approximation techniques, like the k-th order Taylor polynomial technique, are applied and completed by approaches that use least square methods. As will be shown below, such approaches do not conceptually fit this type of non-ergodic model and estimated parameters should remain biased and inefficient. In this document, we suggest a power law (PL)-driven estimation approach, thus moving away from a Gaussian ergodic hypothesis to more general Levy, unstable time (or space) processes, characterized by tail queues, long memory, complex correlation, and plausible convergence to the Gaussian central theorem limit. Once again, as in the case of a labour demand model presented at the end of Part II, the estimation procedure presented could be seen as a generalization to non-ergodic systems from the work of Kullback-Leibler on information divergence (Kullback, 1951 and Golan et al., 1996) on entropy econometrics. Technically, we minimize the Tsallis non-extensive relative entropy criterion function under consistency moment-constraints—incorporating the reparameterized CES function—and regular normality conditions. As such, the approach then encompasses the Bayesian information processing rule while remaining, however, fundamentally based on the second law of thermodynamics[55].

55 For those interested, a vast literature on the subject can be found at: http://polymer.bu.edu/~hes/econophysics/

5.3.1 Power Law and the Constant Elasticity of Substitution (CES) Function

A half century ago, Arrow, Chenery, Minhas, and Solow (1961) proposed a new mathematical function which simultaneously displays the property of homogeneity, constant elasticity of substitution (CES) between factors of production. Additionally, this function presents the possibility of differentiating elasticity of substitution for different industries, sectors, or countries (Klump and Papageorgiou, 2008), thereby generalizing the Cobb-Douglas model. The model was later expanded to other case studies where the system optimally aggregates its components according to some parameters to be specified below. Following Bwanakare (2014), we develop the proposed entropy formalism using a classical CESP explaining the gross domestic product (VA_t) by two classical factors: labour (L_t) and capital (K_t).

The next two cases of the CGE trade model class have been presented in Bwanakare (2014). Let us recall below a CESP mathematical form:

$$VA_t = \alpha\left[\partial K_t^{-\rho} + (1-\partial)L_t^{-\rho}\right]^{-\frac{v}{\rho}} e^{\varepsilon_t} \tag{5.12}$$

or one of its generalized formulations as:

$$Y = \left[\sum_{i=1}^{n} a^{(1-\rho)}{}_i X_i^{\ p}\right]^{-\frac{v}{\rho}} \tag{5.13}$$

where:

$$\rho = \frac{1-\tau^e}{\tau^e} \text{ with } -1 < \rho < +\infty \text{ and } 0 < \tau^e < +\infty \tag{5.14}$$

and

τ^e constant elasticity of substitution between factors, ε_t stands for the random disturbance with unknown distribution. In (5.12), α stands for the shift parameter; the parameter δ belongs to the zero-one interval and represents the share (distribution) of the sold quantities of both distributed factors. Parameter v reflects the degree of changing returns of VA_t to scale. The higher the value of ρ, the higher the degree of substitution between factors. The case of τ^e converging to 0 suggests perfectly substitutable factors. The generalized form (5.13) suggests a case of more than two inputs $X_{i,}$ $i = 1,2,\ldots,n$.

Let us now focus on a useful connection between the CES class of functions and a power law (PL). In fact, to better display that relation, let us aggregate components of model (5.12) into one variable. Then, we get a generic case of a PL of the form:

$$va_t = \alpha\left[\lambda k_t^{-p} + 1\right]^{-\frac{v}{\rho}} e^{\varepsilon_t} = \beta k_t^{\ \eta} e^{\varepsilon_t} \tag{5.15}$$

where, in this case, the endogenous variable va_t is the product per capita. Parameter β represents a general level of technology. The variable k_t stands for a capital coefficient.

The exponent η belongs within the interval (–1, +∞) and defines a per head product elasticity with respect to the capital coefficient. It is evident that the above relation (5.15) can be compared with the relation (2.30) or the achieved tail queue-related Tsallis non-extensive entropy relation (2.31). The random term ε_1, itself, is assumed to follow PL structure. Index *t* means time period. Thus, such PL relationships with the class of CES function seem to engender potential implications on economic grounds, including extensions to the real demand side of the economy. A PL displays interesting properties which may explain its ubiquity at different complexity levels in natural and human organizations. For a survey on the inheritance mechanism and other properties of a PL (Gabaix, 2008). Once again, for the direct relationships between a PL and non-extensive Tsallis entropy, refer to Tsallis (2009). The proposed model generalizes the statistical theory of information approach to non-ergodic systems where q is different to unity. Let us underscore the fact that many findings of recent decades seem to confirm a domination of such systems in the physical real world. Even if we do not intend to set up a philosophical discussion here, a large number of econo-physical contributions of recent years seem to confirm the ubiquity of non-ergodic law in social sciences (e.g., Bottazi et al., 2007, Ikeda et al., 2008, Mantagna et al., 1999). Under such a hypothesis, this could constitute a serious drawback for other competing entropy or econometrics techniques when trying to efficiently model a certain class of phenomena.

The model estimation. We follow the same procedure as in previous examples when we searched to minimize the cross-entropy criterion function under *a priori* moments, including the one concerning the economic model in question. Below, we directly present the model under the reparameterized form of parameters. Then, if we additionally use an escort distribution in moments, the Tsallis cross-entropy econometric model can be stated as:

$$Min \xrightarrow[a,p,,r]{} \alpha\left[\sum a_{jm}\frac{\left[a_{jm}/ao_{jm}\right]^{q-1}-1}{q-1}+\ldots.+\sum p_{km}\frac{\left[p_{km}/po_{km}\right]^{q-1}-1}{q-1}\right]+\beta\sum r_{nj}\frac{\left[r_{nj}/ro_{nj}\right]^{q-1}-1}{q-1} \tag{5.16}$$

s.t.:

$$\ln VA=\ln\left(\sum_{j=1}^{J}g_j\frac{a_j^{\,q}}{\sum_{j=1}^{J}a_j^{\,q}}\right)-\frac{\sum_{h=1}^{H}v_h\frac{w_h^{\,q}}{\sum_{h=1}^{H}w_h^{\,q}}}{\left(\sum_{m=1}^{M}v_m\frac{p_m^{\,q}}{\sum_{m=1}^{M}p_m^{\,q}}\right)}\ln\left[\sum_{i=1}^{I}\left(t_i\frac{b_i^{\,q}}{\sum_{i=1}^{I}b_i^{\,q}}\right)L^{\left(\sum_{m=1}^{M}z_m\frac{p_m^{\,q}}{\sum_{m=1}^{M}p_m^{\,q}}\right)}+\left(1-\sum_{i=1}^{I}\left(t_i\frac{b_i^{\,q}}{\sum_{i=1}^{I}b_i^{\,q}}\right)K^{\left(\sum_{m=1}^{M}z_m\frac{p_m^{\,q}}{\sum_{m=1}^{M}p_m^{\,q}}\right)}\right]+$$

$$+\sum_{n=1}^{N}\sum_{j=1}^{J}z_{nj}\frac{r_j^{\,q}}{\sum_{j=1}^{J}r_j^{\,q}} \tag{5.17}$$

$$\sum\nolimits_{j>2\ldots M}^{J}a_j=1\;\sum\nolimits_{m>2\ldots M}^{M}p_m=1 \tag{5.18}$$

$$\sum_{i>2...I}^{I} b_i = 1 \tag{5.19}$$

$$\sum_{h>2...I}^{H} w_h = 1 \tag{5.20}$$

For reasons of formal presentation, the criterion function (Equation 5.16) does not include probabilities w_h, explaining degree of economy changing to scale and b_i, the parameter of distribution between factors.

In order to improve the estimated parameter quality—in the Bayesian sense—additional constraining data can be added to (5.17–5.20). In the case of a CES model, some economic theory exists. For instance, we can predict sign value domain variation for each parameter. Then we get:

$$0 \leq \alpha = Ga < \infty \tag{5.21}$$

$$-1 \leq \rho = Zp \leq \infty \tag{5.22}$$

$$0 \leq \delta = Tb \leq 1 \tag{5.23}$$

Where α, ρ, δ in Equation 1a stand for the original, "before-reparameterization," parameters. The set G, Z, T stand for the above original parameter support space with the corresponding weight-probabilities a, p, b defining output posteriors. G, Z, T support spaces are defined in the same way as, e.g., Equation 2.45–2.46. Here we just present how we have specified this particular model and not a general rule of specification. Note that depending on error distribution, the weights α, β introduced in the above dual objective function may exercise a significant impact on the model optimal outputs, respectively, as precision and prediction weight. Indeed, the entropy model encompasses statistical losses in the parameter solution space (precision) and in the sample solution space (prediction). As can be easily shown, Lagrange multipliers stand for implicit nonlinear function of the weights (α, β) imposed in the generalized cross-entropy criterion function. Changes in weights alter the corresponding optimal solution value. In general, as in most constrained optimization problems, smaller Lagrange multipliers for a q cross-entropy formulation imply smaller impact of constraints on the objective, in particular for the Tsallis q around unity, i.e., the Gaussian case.

Model outputs. Outputs presented below constitute an important component of the findings in this book. They underscore, more than in previous applications where, generally, outputs from Tsallis entropy fit with those of Shannon entropy, i.e., an illustration of convergence case to Gaussian model. Here, things will change and power law will point out its form. Because of its importance, many details of the model outputs already presented by Bwanakare (2014) will be reported in this section plus new outputs from two additional country case studies. Thus, based on the data source of Table 22, let us present outputs of the three CGE trade models: the production function CESP (already developed through the above sections), the CECS (constant elasticity of commercial substitution function, known as the Armington model), and the CET (constant elasticity of technical transformation). The reasons for presenting these three models of the same class of function are the following: the CES model displays causality relationships; the CECS remains a quasi-identity equation since it is just missing a quasi-constant variable (the indirect taxes) to constitute an identity; and the CET model remains an identity equation, the covariate values of which sum up to the explained value of the model. Statistical data in Table 21 illustrate that situation. Due to a low level of precision, the traditional regression techniques may not be relevant in separating the three cases presented above. We compare the outputs from non-extensive cross-entropy (NCEE) with those from the traditional estimation techniques: the nonlinear least squares (NLS), the generalized methods of moments (GMM), and the maximum likelihood approaches (Green, 2003). Model data first have been dimensioned at logarithmical scale for computational purposes. The computations of the NCEE model were carried out with the GAMS code (General Algebraic Modelling System). Those with the NLS technique were done with Microsoft Excel. Computations by the GMM and ML approaches were executed with a special code from the open source GRETL. Let us first recall the mathematical formulation of the next two CES model classes. A CECS function aggregating interior economic absorption with two business components (locally produced commodities demand (*DO*) and imports (*M*) has the following analytical form:

$$C_t = \alpha\left[\partial DO_t^{-\rho} + (1-\partial)M_t^{-\rho}\right]^{-\frac{1}{\rho}} \tag{5.24}$$

where:

$$\rho = \frac{1-\tau^e}{\tau^e} \text{ with } -1 \prec \rho \prec +\infty \text{ and } 0 \prec \tau^e \prec +\infty$$

and

τ^e constant elasticity of substitution, ε_t stands for random disturbances with unknown distribution.

The last model CET is analytically formulated in the following way:

$$MO_t = \alpha\left[\partial ex_t^{-\rho} + (1-\partial)DMO_t^{-\rho}\right]^{-\frac{1}{\rho}} \tag{5.25}$$

where:

$$\rho = \frac{1-\tau^e}{\tau^e} \text{ with } -\infty \prec \rho \prec -1 \text{and } -\infty \prec \tau^e \prec 0 \tag{5.26}$$

and

ex_t, DMO_t stand for exports and domestically marketed outputs, respectively. The rest of the symbols have the same meaning as in the previous models. The higher the value of ρ, the higher the degree of transformation. When that parameter converges to $-\infty$ we are dealing with a case of perfectly complementary products, which refers to Leontief technology. The case of ρ converging to –1 suggests perfectly substitutable products. In this study, priors were initiated from NLLS outputs. As known, such priors are not deterministically fixed. They are updated according to the Bayesian information processing rule. For simulation purposes, different q-Tsallis parameter values were computed from unity to its admissible[56] highest values minimizing the criterion function. In all models, an *a priori* parameter support space for reparameterization varies between –5.0 and +5.0. The same prior space has been retained for the error disturbance with amplitude varying between –3 and +3, so it conforms to the three sigma rule owing to Chebychev's inequality (Pukelsheim, The Three Sigma Rule, 1994). Both spaces are symmetric around zero. This prevents the estimated parameters from a bias. All the recent works on these subjects (Bwanakare, 2016) seem to confirm that besides the NCEE approach, the NLS remains much better than the remaining econometric methodologies (GMM and ML techniques) for solving this kind of nonlinear inverse problem. Then, in the next model we limit ourselves to the presentation of the model outputs from the NCEE and NLS techniques. Tables 18 and 19 comparatively display the outputs from NCEE and NLS, respectively, for the three models. The NCEE estimator super-consistency for all three models can be noticed despite the small sample. The NLS approach seems better than the GMM procedure as shown in, e.g., Bwanakare (2016). The ML has produced, as theoretically expected, much poorer outputs. Output performance is displayed through error curves in Figures 4 to 9.

5.3.2 Parameter Outputs of the Tsallis Relative Entropy Model

Nonlinear LS estimation outputs. Using traditional nonlinear least square methods, we have linearized the Equations (1a, 1c, and 1d) before applying the Taylor development and the LS approaches.

The NCEE outputs are accurate for all estimated models, and performances of the rest of the econometric approaches seem to be much less competitive. Having used a twelve-year sample in this model, the power law clearly seems to constitute the data generating system. Let us thus comment on these NCEE outputs on the empirical side. The estimated parameters reflect long-run optimal equilibrium values of the system. Since we are dealing with the aggregated accounts of 27 EU countries, the values of estimated parameters seem to reflect our expectations. In particular, in the case of the CESP production model, the estimated parameter p with an estimate

56 In fact, its interval covers Gaussian ($1 \prec q \prec 5/3$) and stable laws (e.g., Levy's) attractors for ($5/3 \prec q \prec 3$).

Table 18: Outputs from the NCEE: dependent var: C(t), MO(t), VA(t).

Exogenous var:	***A***	***δ***	***p***	***v***	***I(Pr)***	***CV***
CESP(L(t), K(t))	1.866	0.163	0.001	1.000	0.99	0.006
CET(DO(t),Ex(t))	2.000	0.5	-1.0001		0.999	4.271E-7
CECS(C(t),M(t))	2.000	0.499	pu -0.985		0.999	2.705E-5

q Tsallis parameter (weight α_i=0.05) = 2.333 (CECP)
q Tsallis parameter (weight α_i=0.05) = 1.0001 (CECS)
q Tsallis parameter (weight α_i=0.05) = 1.0001 (CET)

Table 19: Outputs from the NLS, models CESP, CET, CECS: dependent var: C(t), MO(t), VA(t)

Exogenous var:	***A***	***δ***	***p***	***v***	***R^2***
CESP(L(t), K(t))	1.995	0.282	3.046	0.993	
Parameters T-value)	48.89	6.61	1.49	6.61	0.88
CET(DO(t),Ex(t))	2.008	0.497	-0.954		
Parameters T-value	558.120	1551.353	-292.532		0.999
CECS (C(t), M(t))	2.147	0.477	-0.532		
Parameters T-value	6.257	13.073	-1.554		0.83

Table 20: Outputs from the NCEE: dependent var: GDP(t)

COUNTRY	***A***	***δ***	***p***	***Q***	***EC***	**n**
Belgium	2.326	0.15	- 3.524398E-7	7/3	0.027	18
USA_sic33	0.777	0.057	-9.84248E-5	7/3	3.373000E-4	20

around zero suggests a convergence of the analysed function to the classical *Cobb-Douglas function*, displaying in the present case constant returns to scale. A long-run optimal equilibrium share parameter δ between factors shows a lower proportion of labour of around 16.3% with respect to capital share (83.7%). In 2010, this proportion was around 57% for labour. Thus, according to these outputs, long-run optimal production will require much less labour demand, around 16%. For the models CET and CECS, the estimated parameters show, in the long-run, a quasi-perfect substitutability and a balanced share between local and foreign commodities. Expected free trade

Table 21: Outputs from the NLS: dependent var: GDP(t)

Exogenous var:	***A***	***δ***	***p***	***v***	R^2
Belgium	11,85713	1,82255	2,833935	0,867049	0.999
USA_sic33	0,00038	4,85717	0,061815	1,054080	0.999

Table 22: Aggregated data (chain-linked volumes at 2005 exchange rates) for models (in 1000 billion euro).

year	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
VA_mld	7,82	8,136	8,427	8,578	8,978	9,294	9,76	10,288	10,288	9,777	10,149	10,412
K_mld	3,287	3,42	3,55	3,641	3,86	4,02	4,265	4,528	4,476	4,105	4,336	4,444
L_mld	4,427	4,606	4,76	4,819	4,99	5,149	5,374	5,63	5,683	5,554	5,691	5,834
MO	8,691	9,026	9,212	9,325	9,442	9,689	9,877	10,211	10,550	10,621	10,162	10,359
DMO	5,719	5,693	5,751	5,792	5,846	5,813	5,773	5,711	5,795	5,797	5,924	5,693
export	2,971	3,333	3,461	3,533	3,596	3,876	4,104	4,500	4,754	4,825	4,238	4,666
Imports	2,938	3,272	3,361	3,414	3,526	3,796	4,029	4,411	4,672	4,727	4,153	4,551

Source: *http://appsso.eurostat.ec.europa.eu/nui/setupModifyTableLayout.do*

barriers and a closer level of productivity among world business partners could advocate in favour of such outputs. For simulation purposes, different Tsallis-q parameter values were computed over the domain of definition of $q \in [1, 3]$ which covers the Gaussian basin of attraction $[1 < q < 5/3]$ and Levy's attractors law $[5/3 < q < 3]$. The q-parameter has been incremented by a step of 0.25 starting from unity (the Shannon entropy point). These different values of q generated the model error disturbances which allowed the computation of the error coefficient variation (CV). The index CV is obtained by dividing the model standard error by the average value of the dependent variable. All the above models present an error coefficient variation (CV) of around zero. The Tsallis Information Index (Bwanakare, 2014) presented in Part II of this book is around unity for the three models, suggesting relatively close to zero information divergence between priors and posteriors, under given model restrictions. We would have expected optimal solutions for q less than 5/3 or, in the worst case, less than 2 for theoretical and empirical evidence. This is the case for the two commercial CET and CECS models where q is almost equal to unity, suggesting a Gaussian distribution. For the CESP production model, minimum LS errors are obtained for q

Figure 4: Error term for NLS, cross-entropy and GMM estimated models (CECS model).

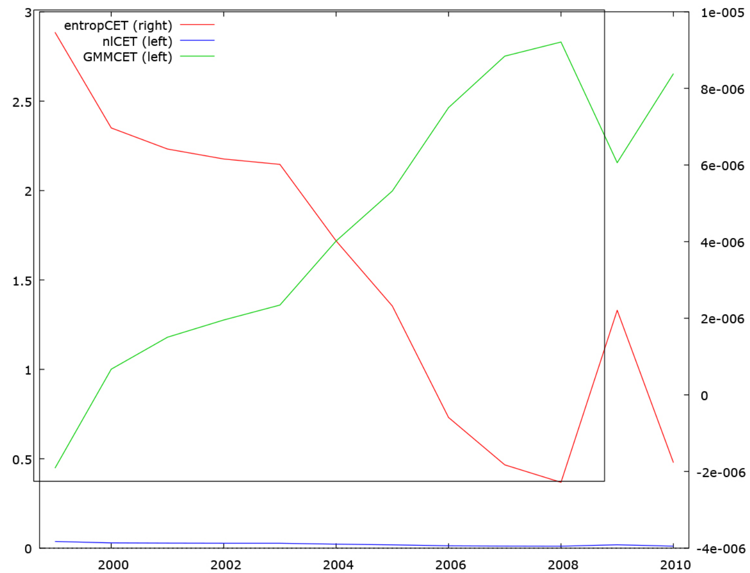

Figure 5: Error term for NLS, cross-entropy and GMM estimated models (CET model).

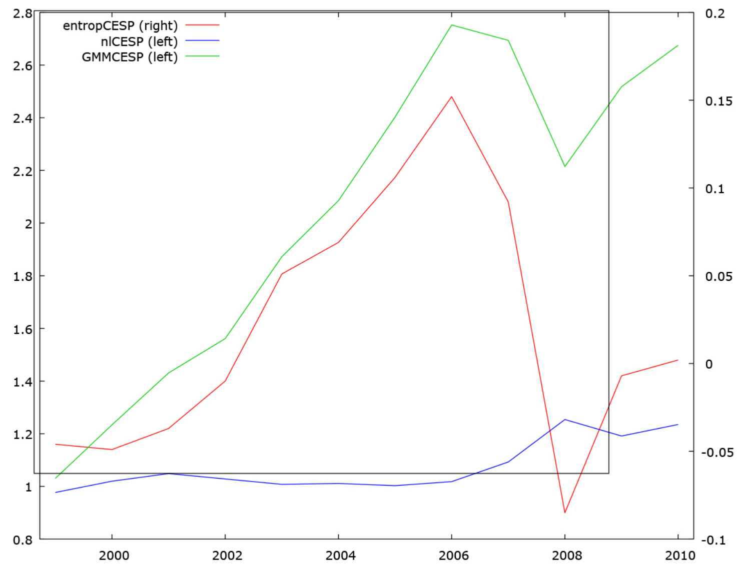

Figure 6: Error term for NLS, cross-entropy and GMM estimated models (CESP model).

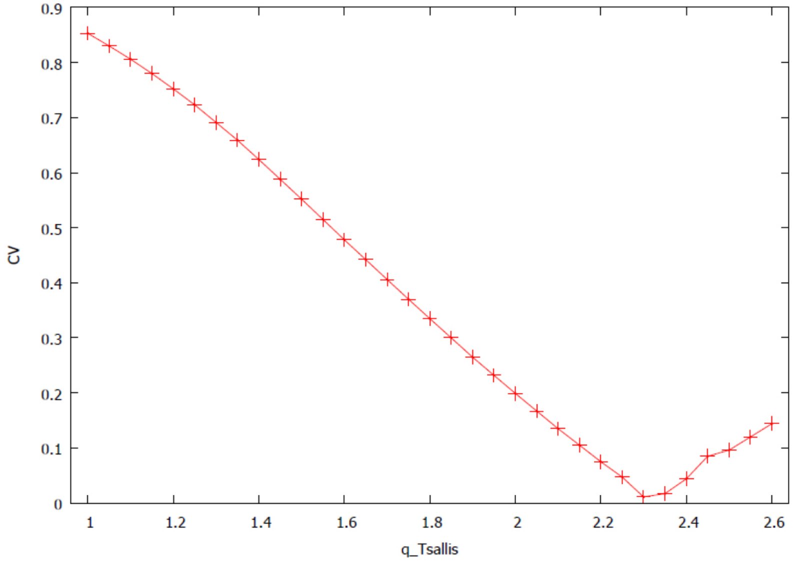

Figure 7: Model disturbance (CV) curve as a function of q, for [1<q<2.6] (model CESP).

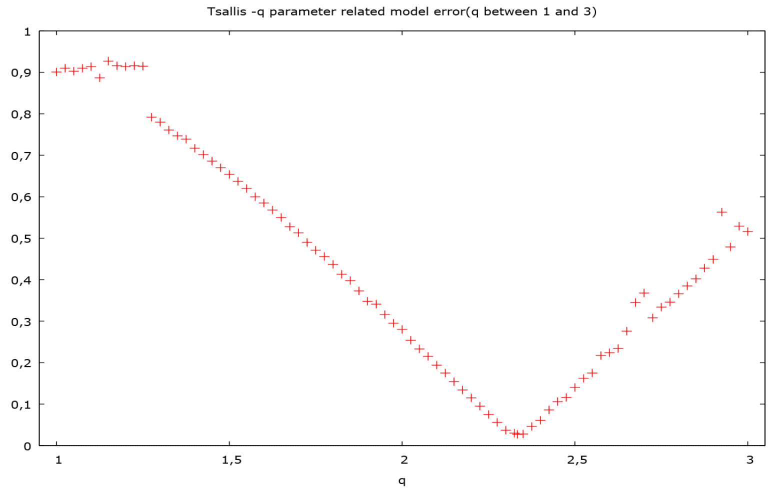

Figure 8: The Tsallis q-parameter related model error for the Belgium model (q between 1 and 3).

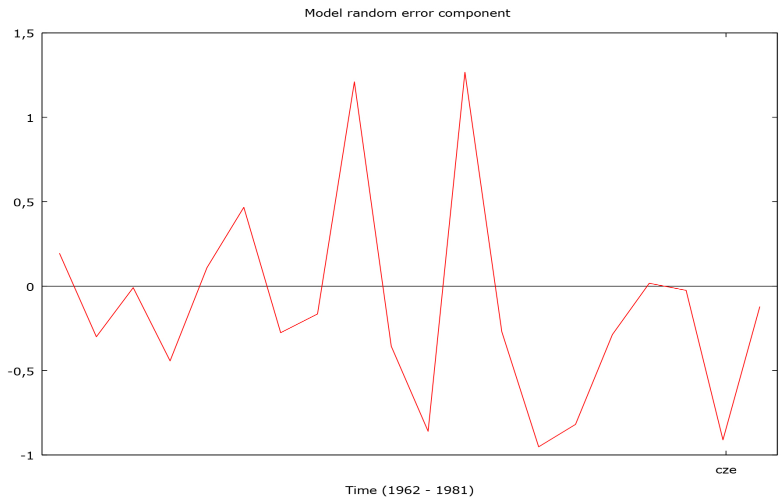

Figure 9: Model random error coefficient for SIC 33: the USA Primary Metals production NCEE model.

around 7/3 (see Figure 6). However, comparable outputs from Borges (2004) exist. He found cumulative distribution of the scaled gross domestic product of 167 countries around the world for the year 2000 corresponding to q = 3.5. Figure 7 and 8 display a convex space defining different optimal CV values owing to different simulated q-Tsallis parameters for the CESP model. Minimum CV corresponds to the minimum of information divergence or of the sum of geometrical error of least squares. To verify the Tsallis related model outputs, we have computed a classical S-K-L cross-entropy econometric model which has produced for all the three models, as expected, the same values as those obtained from Tsallis formalism for *q* equal to unity. Such trivial results have not been reported in the above output tables. In the case of the CESP model, we found for *q* converging to unity a CV of 85.3% in the case of K-L, against 0.06% for an optimal *q* equal to 2.331. Thus, this point shows the advantage of modelling with non-extensive entropy rather than Shannon entropy, reduced to *q* equal to unity. Finally, in Tables 20 and 21 model outputs are limited to two CESP outputs of Belgium and the USA Primary Metals production model. Contrary to all the models that have been presented in this section, this last USA study treats a space model, already discussed in Green[57]. As expected, the outputs from the NCEE reveal a higher precision than those from the NLS approach. In spite of a high coefficient of determination displayed R^2, economic theory related signs and space area of parameters remain inappropriate in the case of NLS outputs.

5.4 Conclusions

The present example has presented the rationale of the proposed Tsallis cross-entropy approach for unstable, nonlinear econometric models in a more elegant way than in the preceding applications. Though the experiment was limited to three different CES models with respect to their stochastic forms, a large class of economic and financial models could fall into this category. Only outputs produced by Tsallis formalism reflect these stochastic differences *a priori* known. The Tsallis entropy super-convergence estimator should only be explained, even in this unique but complex experiment, by the data generating Pl distribution. More investigation, particularly on the ARFIMA class of models, is needed to confirm the above findings and the importance of the PL approach in econometrical modelling.

Bibliography – Part V

Adelman, I., & Robinson, S. (1978). *Income Distribution Policy in Developing Countries: A Case Study of Korea*. Stanford: Stanford University Press.

Armington, P.S. (1969). *A theory of demand for products distinguished by place of production*. IMF Staff papers 16, pp. 159–76.

Arndt C., Tarp, F., Robinson, S., Jensen, H.T., & Cruz, A. (2001). Parameter estimation for a Computable general equilibrium model: A maximum entropy approach. Washington: Working Paper, International Food Policy Research Institute.

57 See Green (2003).

Arndt C., Finn, T., Sherman, R., Jensen, H.T., & Cruz, A. (1998). Social Accounting Matrices for Mozambique 1994 and 1995. Washington, D.C. 20006 U.S.A.: Working Paper, International Food Policy Research Institute.

Arrow, K.J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. Econometrica, 22, 265–90.

Arrow, K.J., & Hahn, F.H. (1971*). General Competitive Analysis*. San Francisco: Holden Day.

Arrow, K.J., Chenery, H.B., Minhas, B.S., & Solow, R.M. (1961). Capital-Labor Substitution and Economic Efficiency. *The Review of Economics and Statistics*, 43(3), pp. 225–250.

Bento, A.M., & Jacobsen, M. (2007). Ricardian rents, environmental policy and the 'double-dividend' hypothesis. *J. Environ. Econ. Manag.*, 53 (1), pp. 17–31.

Bottazzi G., Cefis, E., Dosi, G., & Secchi, A. (2007). Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics*, 29(1), pp. 137–159.

Bourguignon, F., Michel, G., & Miqueu, D. (1983). Short Run Rigidities and Long Run Adjustments in a Computable General Equilibrium Model of Income Distribution and Development. *Journal of Development Economics*, pp. 1–2.

Bwanakare, S. (2014). Non-Extensive Entropy Econometrics: New Statistical Features of Constant Elasticity of Substitution-Related Models. *Entropy*, pp. 2713–2728.

Bwanakare, S. (2016). Non-Extensive Entropy Econometrics and CES production Models: Country Case Study. Journal of the International Association for Official Statistics (IAOS). Retrieved from: http://content.iospress.com/articles/statistical-journal-of-the-iaos/sji1021

Capros, P., Karadeloglou, P., & Mentzas, G. (1990). An empirical assessment of macroeconometric and CGE approaches in policy modeling. *Journal of Policy Modelling*, 12 (3), pp. 557–585. doi:http://dx.doi.org/10.1016/0161-8938(90)90013

De Melo, J. (1988). Computable General Equilibrium Models for Trade Policy analysis in Developing Countries: A Survey. *Journal of Policy Modeling*, 4, pp. 469–503.

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

Decaluwé, B., Martens, A., & Savard, L. (2001). *La politique économique du développement et les modèles d'équilibre général calculable*. Montréal, Canada: Les Presses de l'Université de Montréal.

Devarajan, S., & Robinson, S. (1983). *Selected Bibliography on Computable General Equilibrium Models of Developing Countries*. Harvard and U.C. Berkeley: Mimeo.

McDonald, S., Thierfelder, K., Robinson, S. (2007). *Globe: A SAM Based Global CGE Model using GTAP Data*. Retrieved from www.usna.edu/EconDept/RePEc/usn/wp/usnawp14.pdf

Evans, G.W., & Ramey, G. (2006). Adaptive Expectations, Underparameterization and the Lucas Critique. *Journal of Monetary Economics*, 53, pp. 249–264.

Francois, J. (2001). *Flexible Estimation and Inference Within General Equilibrium Systems*. Discussion Paper. 0129. Adelaide 5005, Australia: Adelaide University.

Frasera, L., & Waschik, R. (2013). The Double Dividend hypothesis in a CGE model: Specific factors and the carbon base. *Energy Economics*, 39, pp. 283–295.

Gabaix, X. (2008). *Power laws in economics and finance*. Retrieved from: http://www.nber.org/papers/w14299

Giffin, A. (2009). From Physics to Economics: An Econometric Example Using Maximum Relative Entropy. *Physica A: Statistical Mechanics and its Applications*, pp. 1610–1620.

Ginsburg, V., & Waelbroeck, J. (1981). *Activity Analysis and General Equilibrium Modelling*, Amsterdam North-Holland. Amsterdam: North-Holland.

Ginsburgh, V., & Keyzer, M. (1997). *Structure of Applied General Equilibrium Models*, MIT Press. MIT Press.

Ginsburgh, V., & Waelbroeck, J. (1976). *Computational Experience with a Large General Equilibrium Model in computing equilibria*. Amsterdam North-Holland.
Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2014). Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach. Retrieved from https://www.gtap.agecon.purdue.edu/resources/download/7136.pdf
Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015). Estimating parameters and structural change in CGE models using a Bayesian cross-entropy estimation approach. Policy Research working paper no. WPS 7174.
Golan, A., Karp, L.S., & Perloff, J.M. (1996). Esimating a mixed strategy employing maximum entropy. Working paper, California Agricultural Experiment Station. California.
Golan, A., Judge, G., & Miller, D. (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*, Chichester, England: Wiley.
Gossen, H.H. (1983). *The laws of human relations and the rules of human action derived therefrom*. Cambridge: MIT Press.
Green, W. (2003). *Basics of Econometrics*. NY: Hall Prencite, 5th edition.
Guerrien, B. (2000). *Dictionnaire d'analyse économique*. Paris, France: La Decouverte.
Harberger, A.C. (1959). The Corporation Income Tax: An Empirical Appraisal. *Tax Revision Compendium* 1, pp. 231–240.
Harberger, A.C. (1962). The Incidence of the Corporation Income Tax. *Journal of the Political Economy*. 70, pp. 215–240.
Harris, R., & Cox, D. (1983). Trade, Industrial Policy and Canadian manufacturing. Canadian journal of economics.
Herault, N. (2006, March). Building And Linking A Microsimulation Model To A CGE Model For South Africa. *South African Journal of Economics*, 74(1), pp. 34–58.
Ikeda, Y., & Souma, W. (2008). International Comparison of Labour Productivity Distribution. (Cornell University Library) Retrieved from arXiv:0812.0208v4
Jaynes, E.T. (1988). *Maximum-entropy and Bayesian methods in Science and Engineering, Vol. One: Foundations*. Boston, USA: G.J. Erickson and C.R. Smith, Kluwers Accademic Publishers.
Jevons, W.S. (1879). *The Theory of Political Economy*. London: Macmillan.
Jian, X. (2000). An Environmentally Extended Social Accounting Matrix, *Environmental and Resource Economics*. 16(4), pp. 391–406.
Jorgenson, D.W. (1984). Econometric Methods for General Equilibrium Analysis In *Applied General Equilibrium Analysis*. Cambridge: H. Scarf and J. Shoven, Eds., Cambridge University Press.
Jorgenson, D.W. (1998a). *Growth. Volume 1: Econometric General Equilibrium Modeling*. Cambridge: The MIT Press.
Judge, G.G. & Mittelhammer, R.C. (2012). *An Information Theoretic Approach to Econometrics*. Cambridge, UK: Cambridge University Press.
Klump, R., & Papageorgiou, C. (2008). The CES Production Function in the Theory and Empirics of Economic Growth. *Journal of Macroeconomics*, 2, pp. 599–600.
Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, 22, 79–86.
Kydland, F.E., & Prescott, E.C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica*, 50 (6), pp. 1345–1370. doi:10.2307/1913386. JSTOR 1913386
Kydland, F., & Prescott, E.C. (1977). Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, pp. 473–492. doi:10.1086/260580
Leontief, W. (1941). *The Structure of American Economy, 1919–1929*. Cambridge: Harvard University Press.
Lofgren, H., Harris, R.L., & Robinson, S. (2000*). A Standard Computable General Equilibrium (CGE) Model in GAMS*. Washington: Intl Food Policy Res Inst.

Lucas, R. (1976). Econometric Policy Evaluation: A Critique. 1, 19–46. doi:10.1016/S0167–2231(76)80003–6
Mansur, A., & Whalley, J. (1984). Numerical Specification of Applied General Equilibrium Models: Estimation, Calibration, and Data in Scarf and Shoven. *Applying General Equilibrium Analysis*, pp. 69–127.
Mantegna, R.N., & Stanley, H.E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
Marshall, A. (1890). *Principles of Economics*. London: Macmillan.
McKenzie, L.W. (1959). On the Existence of General Equilibrium for a Competitive Economy. *Econometrica*, 1, pp. 54–71. doi:10.2307/1907777. JSTOR 1907777
McKenzie, L.W. (1981). The Classical Theorem on Existence of Competitive Equilibrium. *Econometrica*, 4, pp. 819–841. doi:10.2307/1912505. JSTOR 1912505
McKenzie, L.W. (1987). Turnpike theory. *The New Palgrave: A Dictionary of Economics*, pp. 712–720.
McKenzie, L.W. (1954). On Equilibrium in Graham's Model of World Trade and Other Competitive Systems. *Econometrica*, 2, pp. 147–161.
Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, 3, pp. 315–335. doi:10.2307/1909635
Nash, J.F. (1950). The Bargaining Problem. *Econometrica*, 2, pp. 155–162. doi:10.2307/1907266. JSTOR1907266
Pearce, D. (1991). The Role of Carbon Taxes in Adjusting to Global Warming. *The Economic Journal* (101), pp. 938–948.
Pukelsheim, F. (1994). The Three Sigma Rule. *American Statistical Association*, 2, pp. 88–91.
Pyatt, G., & Thorbecke, E. (1976). *Planning Techniques for a Better Future*. Geneva, Switzerland: International Labour Office.
Pyatt, G. (1988). A SAM approach to modeling. Elsevier, *Journal of Policy Modeling*, 10(3), pp. 327–352.
Pyatt, Graham, & Roe, A. (1977). *Social Accounting Matrices for Development Planning: With Special Reference to Sri Lanka*. Cambridge: Cambridge University Press.
Sasmaz, M.U. (2016). Validity of double dividend hypothesis in EU–15 countries. *Global Journal on Humanities & Social Sciences*, pp. 30–36.
Scarf, H. (1984). *The Computation of Equilibrium Prices*. Cambridge: (H. Scarf and J. Shoven, Eds.), Cambridge University Press.
Scarf, H. (1969). An Example of an Algorithm for Calculating Equilibrium Prices. *American Economic Review*, 59.
Shoven, J.B., & Whalley, J. (1972). A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in *U.S. Journal of Public Economics* 1 (3–4), pp. 281–321.
Shoven, J.B., & Whalley, J. (1992). *Applying General Equilibrium*. New York: Cambridge University Press.
Sims, C.A. (1987). A Rational Expectations Framework for Short-Run Policy Analysis. *New Approaches to Monetary Economics*, pp. 293–308.
Sims, C.A., 1987. A rational expectations framework for short-run policy analysis. In: Barnett, W.A., Singleton, K.J. (Eds.), *New approaches* to *monetary economics*. Cambridge University Press, Cambridge, UK, pp. 293–308.
Sims, C.A. (2007). *Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian*. Princeton : Princeton University.
Stone, R. (1955). *Input-Output and the Social Accounts*. New York: J. Wiley.
Stone, R., & Brown, A. (1962). *A computable model for economic growth*. Cambridge, UK: Cambridge Growth Project.

Taheripour, F., Khanna, M., & Nelson, C.H. (2008). Welfare impacts of alternative public policies for agricultural pollution control in an open economy: a general equilibrium framework. *Am. J. Agric. Econ*., 90(3), pp. 701–718.
Takeda, S. (2007). The double dividend from carbon regulations in Japan. *J. Jpn. Int. Econ*., 21, pp. 336–364.
Tovar, C. (2009). DSGE Models and Central Banks. Economics, 3 (2009–16)(1). doi:10.5018/economics-ejournal.ja.2009–16
Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. Berlin: Springer.
Walras, L. (1874). Éléments d'économie politique pure, ou, Théorie de la richesse sociale. Lausanne: L. Corbaz & Cie.

Part VI: **From Equilibrium to Real World Disequilibrium: An Environmental Model**

1 Introduction

Through the use of adequate mathematical tools, recent works have attempted to better adapt the theoretical general equilibrium framework to realities examined by different schools of thought in economics where conditions of imperfect competition are taken into account. In many situations, a benchmark device allowing for assessing deviation from the theoretical optimal equilibrium is the concept of market efficiency in the Pareto context[58]. Standard macroeconomic theory holds that the conditions leading to disequilibrium are market failure, market imperfections, the search for non-economic targets, and international business. Let us briefly describe these four factors. *Market failure* takes place in the presence of public goods or externalities in production or consumption or when some economic agents are in possession of imperfect information. The common attribute of these market anomalies is their ability to appear even in the presence of perfect competition over all markets. When these market distortions are not removed, the economy moves from a socially optimal trajectory of welfare in the sense of Pareto. Externalities, as a cause of market failure, will be examined later in the context of the environmental economy.

Imperfect markets are related to the supply-and-demand side of commodity and factor markets. It is generally admitted that the level of market concentration of a given business is a good indicator of market imperfection. In the case of commodity supply, imperfect competition will shift down production and consumer utility. In the case of factor markets, imperfections lead to lost productivity.

As far as *searching for non-economic targets* is concerned, the role of government, usually deemed to help in correcting market failures or imperfections, can have a distorting role. Thus, government will search for non-economic targets[59] of various attributes like ethical values, political agenda of politicians in connection with the political business cycle, and lobbying interests with the problem of rent seeking. The common attribute of these non-economic targets is that they lead to social efficiency losses.

International business policy may lead to economic inefficiency on the side of commodity or production factor. Thus, one can point to, for instance, the impact of fiscal or monetary policies as regulators of relative prices and market competition. These policies would thus create market distortion and lead to disequilibrium.

58 However, J.M. Buchanan (Nobel Prize, 1986) presents a controversial point of view on the subject. See, e.g., Alain Marciano, "Why Markets Do Not Fail"; Buchanan on voluntary cooperation and externalities (Nov. 2010); or A.H. Barnett and Bruce Yandle, "The End of the Externality Revolution" (2005).
59 For this vast subject see: Downs, Economics Theory of Democracy (1957); T. Buchanam, The Theory of Public Choice (1972); W.D. Nordhaus, political business cycle (1975); Breton, A. and Wintrobe, R., The Logic of Bureaucratic Conduct: An Economic Analysis of Competition, Exchange, and Efficiency in Private and Public Organizations (1982); W.A. Niskanen, Bureaucracy and Public Economics (1971).

The next section presents the case of pollution as an externality exemplifying market failure in the economy. Later, we will balance an environmentally extended and unbalanced social accounting matrix. Let us first show, below, the impact of pollution on economic equilibrium.

Economic efficiency, perfect competition, and externalities. To show how externality generated by private activity disrupts the economy from equilibrium, let us take the case of two producers, the first generating pollution (for instance, industry) and the second being affected by it (for instance, agriculture).

First, let us rewrite below the definition of the marginal rate of transformation of products of the economy:

$$TmTP = \frac{\frac{\delta q_2}{\delta x_2}}{\frac{\delta q_1}{\delta x_1}} = \frac{Cm_1}{Cm_2}$$

where q_j explains a quantity produced by producer j (j=1,2), and cm_j the marginal cost for each producer *j*.

Now, suppose producer 2 (industry) generates a negative externality that affects producer 1 (agriculture).

In this case, production function of producer 1 has to be rewritten as:

$$q_1 = f_1(x_1, q_2)$$

The marginal physical productivity of the input used by producer 1 is, therefore:

$$\frac{\delta q_1}{\delta x_1} + \frac{\delta q_1}{\delta q_2}\frac{\delta q_2}{\delta x_2}$$

The second side of the formula above represents externality and has, for this case, a negative value, since $\delta q_1 / \delta q_2 \prec 0$. Thus, from the above definition of *TmTP*, we see that:

$$Cm_1(E) \succ Cm_1$$

This shows that the negative externality from producer 2 increases the marginal cost of producer 1 to a new, higher level, that is:

$$TmTP(E) = \frac{\frac{\delta q_2}{\delta x_2}}{\frac{\delta q_1}{\delta x_1} + \frac{\delta q_1}{\delta q_2}\frac{\delta q_2}{\delta x_2}} = \frac{Cm_1(E)}{Cm_2}$$

with $TmTP(E) \succ TmTP$. We note that when $\delta q_1 / \delta q_2 \succ 0$, we then have positive externalities (e.g., public goods). This case will not be dealt with here.

In economic terms, if there is a negative externality on good 1 from good 2, for an increase in the production of unit 1 good, at the level of the total economy, we must accept losing more production of good 2. Thus, optimal conditions are no longer valid.

2 Extending to an Environmental Model

Let us start by giving hypotheses for the model. The first hypothesis states that pollution is treated as a collective good associated with private production. This means it enters inside the utility function of all agents. The second hypothesis has pollution being treated in the model as a factor of production since it can be substituted for classical inputs through pollution abatement activities as alternative uses. Thus abatement activities might be regarded as an opportunity cost for productive inputs in the context of general equilibrium. The third and last hypothesis is that the model is of general equilibrium.

The model relations. The objective is to maximize utility of representative consumer 1, that is:

$U_1(X_{i1}, E)$

under constraints:

a) the utility of each consumer j other than 1 is predetermined and is at least equal to U_{1A}; that is:

   $U_j(X_{ij}, E) \geq U_{jA}$
   i=1…n
   J=2…m.

b) production functions of all enterprises k are already predefined, i.e.,

   $Y_k = F_k(L_{ik}, E_k) \geq U_{jA}$
   K=1…h,
   i=1…n.

c) the general equilibrium constraint is:

   $$\sum_{1}^{m} X_{ij} - \sum_{k}^{h} L_{ik} \leq R_i$$

   where:
   L_{ik}: inputs involved in production by enterprises,
   X_{ij}: consumed goods,
   R_i: available resources inside the economy,
   E_k: externalities or pollution as an input.

To solve for the above model, we need to formulate its Lagrange as follows:

$$V = U_1(X_{1i}, E) + \sum_{j=2}^{m} \lambda_j \left[U_j(X_{ij}, E) - U_{jA}\right] + \sum_{k=1}^{h} \mu_k \left[Y_k - F_k(L_{ik}, E_k)\right] +$$

$$+ \sum_{i=1}^{n} \omega_i \left(R_i - \sum_{j=1}^{m} X_{ij} + \sum_{k=1}^{h} L_{ik} \right)$$

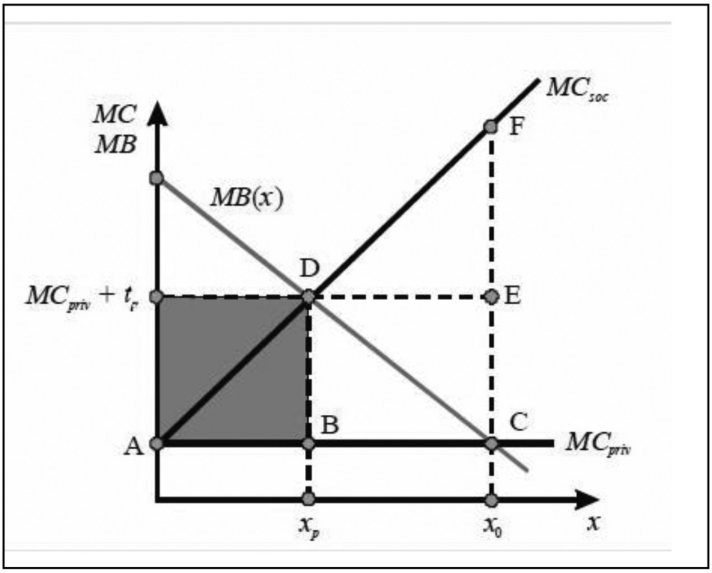

Figure 10: Optimal Pigovian tax

where α_j, μ_k and ω_i are Lagrange multipliers.

Deriving the model for the first order conditions with respect to independent variables L_{ik} X_{ij} and E_k, and checking for the second order, we finally obtain the following two optimum conditions:

1. $\mu_k . \frac{\partial F_k}{\partial E_k} = \frac{\partial U_1}{\partial E} + \sum_{j=2}^{m} \lambda_j . \frac{\partial U_j}{\partial E}$

2. $\lambda_j . \frac{\partial U_j}{\partial X_{ij}} = \mu_k . \frac{\partial F_k}{\partial L_{ik}}$

The first condition says that firms will carry out production activity up to the level where marginal productivity of pollutant emission (lhs of cond. 1) equals marginal cost of that emission (rhs of cond. 1), i.e., the social loss of utilities associated with production.

The second condition means that weighted marginal utilities of consumers *j* must equalize weighted marginal productivity of firms k, for any consumed good X_i or used input L_i.

Criterion of externality internalization. The optimum presented above represents the social and economic optimum as defined by equalization of marginal profit of producers (mpp) with social marginal cost (smc) of pollution owing to production activity. The principle polluter-payer results when smc ≥ mpp. The next figure displays the essence of the above formula.

Without environmental regulations, the market equilibrium is x_o. As shown in Figure 10, the welfare loss in such equilibrium is represented by the box CDF, where the marginal social costs (MC_{soc}) exceed the marginal benefits of consumption (MB(x)) for all units consumed in excess of x_p. Tx is a tax which is equal to the marginal environmental damage (MED). It would make recovery of the social damages CDF possible. It is represented by a grey area.

Before ending this section, it is worthwhile to add that, apart from taxes, there are many other environmental policy instruments such as efficiency targets, quotas, and tradable permits.

2.1 The Model of Carbon Tax

The following models include carbon tax policy within sector activity and institutions and show the uniqueness of the solution in the context of the much revisited CGE model.

Production of goods:
Production function adopted here is the one taken from Cobb-Douglas technology for the production of two goods (good 1 and good 2) with capital, labour, and the environmental pollution input poll.

$$x_i = An^{\alpha_i} k^{1-\alpha_i} e^{\gamma \cdot POLLtax}$$

Tax on Pollution *POLLtax*
Let us define pollution from the production in the following way:

$$POLLtax_i = E_{tax} \cdot x_i$$

E_{tax}: tax on pollution per one unit, one unit of goods used or consumed.

Demand for production factors:
The cost function with tax can be presented as follows:

$$\phi_i(w, r, x_i) = wn_i + rk_i + TAX_{POLL} \cdot P_i \cdot x_i$$

w = return to labour, r = return to capital, and n_i and k_i the amount of labour and capital, respectively. TAX_{POLL} is the tax rate per unit of pollution. Given this level of return, optimal inputs that minimize the above function for the production of x_i become:

$$n_i = A_i^{-1}\left(\frac{\alpha_i r}{(1-\alpha_i)w}\right)^{1-\alpha_i} \cdot x_i$$

$$k_i = A_i^{-1}\left(\frac{\alpha_i r}{(1-\alpha_i)w}\right)^{-\alpha_i} \cdot x_i$$

Supply of goods:
With perfect competition, the price of a good equals the marginal cost of production:

$$p_i = \frac{\partial\phi_i(w,r,x_i)}{\partial x_i} = A_i^{-1}\alpha_i^{-\alpha_i}(1-\alpha_i)^{(\alpha_i-1)} w^{\alpha_i} r^{(1-\alpha_i)} + TAX_{POLL} \cdot P_i$$

Households:
The income of households is defined as expected:

$m = wn + rk + taxtransfer$

and the budget restriction is:

$m = p_1 c_1 + p_2 c_2$

Utility function:
Let us assume the household/households to have the classical Cobb-Douglas utility function but where pollution is added as a separate term; that is:

$$U(H) = c_1^{\beta} \cdot c_2^{1-\beta} \cdot e^{\gamma \cdot POLL}$$

β states how the households value their consumption of good 1 and good 2.
γ is the parameter that states how the households value pollution. If γ is negative the households are assumed to dislike pollution.

Demand:
When we combine the budget restriction and the utility function of the household, we can derive the demand function for the different goods:

$$c_1(p_1, p_2, m) = \frac{\beta m}{p_1}$$

and,

$$c_2(p_1, p_2, m) = \frac{(1-\beta)m}{p_2}$$

Equilibrium conditions:
We have two equilibriums to be formulated below: one concerns the goods market and prices and the other the factor market.
Goods market:

The supply (i.e., the production) should equal the demand (i.e., the consumption of households):

$$x_1 = \frac{\beta m}{p_1}$$

$$x_2 = \frac{(1-\beta)m}{p_2}$$

Ceteris paribus, the price equals marginal cost of production:

$$p_1 = A_1^{-1}\alpha_1^{-\alpha_i}(1-\alpha_1)^{(\alpha_1-1)}w^{\alpha_1}r^{(1-\alpha_1)} + TAX_{POLL}\cdot P_1$$

$$p_2 = A_2^{-1}\alpha_2^{-\alpha_i}(1-\alpha_2)^{(\alpha_2-1)}w^{\alpha_2}r^{(1-\alpha_2)} + TAX_{POLL}\cdot P_2$$

Factor market:

The supply (i.e., the households' total resources) should equal the demand for the different factors, and then we have an optimal factor market:

$$n = A_1^{-1}\left(\frac{\alpha_1 r}{(1-\alpha_1)w}\right)^{1-\alpha_1} x_1 + A_2^{-1}\left(\frac{\alpha_2 r}{(1-\alpha_2)w}\right)^{1-\alpha_2} x_2$$

$$k = A_1^{-1}\left(\frac{\alpha_1 r}{(1-\alpha_1)w}\right)^{\alpha_1} x_1 + A_2^{-1}\left(\frac{\alpha_2 r}{(1-\alpha_2)w}\right)^{\alpha_2} x_2$$

We then have six equations and six endogenously determined variables x1, x2, p1, p2, w and r. This means that such an extended CGE model has a solution.

2.2 Carbon Tax Model and Double-Dividend Hypothesis

Pigou (1932) suggested imposing a tax per unit of pollution at a rate t_α, equal to the marginal external damages per unit of pollution. Carbon tax means that not only the level of taxes depends upon the consumed quantity of carbon but also on its consumed quality. The targeted objective being the reduction of more polluting energy (e.g., that which results in CO_2) in favour of cleaner energy technologies. At the same time, such a strategy allows for an interiorizing of negative externalities through introduction of an ad hoc tax that is socially equitable. Accordingly, such a tax policy should help in recovering lost macroeconomic equilibrium and social Pareto optimum welfare under the free market hypothesis. However, like any tax, a carbon tax should negatively impact economic growth.

Thus, this leads to a fundamental question as to whether or not the above mentioned positive effects are sufficient to explain its introduction. Tullock (1967) raised the possibility that government revenue would be "free," while Terkla (1984) estimated the amount of revenue and the efficiency gains from using it. For Lee and Misiolek (1986), the whole benefit of imposing a pollution tax depends on whether it raises revenue. We now show how this emphasis on revenue continues in the double-dividend literature.

David Pearce (1991) is probably the first writer to use the term *double-dividend*:

> Governments may then adopt a fiscally neutral stance on the carbon tax, using revenues to finance reductions in incentive-distorting taxes such as income tax or corporation tax. This "double-dividend" feature of a pollution tax is of critical importance in the political debate about the means of securing a "carbon convention."

Thus, the idea is that budgetary income from a carbon tax will be used to reduce other taxes so that one will get positive effects not only with respect to environmental targets but also economic targets.

Validity of the hypothesis (strong form)

The validity of the double-dividend hypothesis cannot logically be settled as a general matter. For instance, the environmental tax may have its own distorting effects on labour supply and, therefore, can have the same excess burden as a tax on labour income. Each proposal must be evaluated individually. This evaluation must fully specify the policies already in place as well as the reform under consideration.

3 Compensatory and Equivalent Variations: Two Types of Welfare Measurement

The next two questions will help us understand the difference between compensatory and equivalent variations: If relative price of two given commodities had to change, what would the change in income be if we needed to maintain the consumer at the same welfare level given this new situation? If relative price did not change, what would the equivalent change in income be that could produce the same impact as the relative price?

A) **Compensatory and equivalent variations.** Let us adopt the Cobb-Douglas behavioural function and define a utility function as follows:

$$U = q_1^{\gamma} q_2^{1-\gamma}$$

U: level of direct utility;
q_1: consumed quantity of good 1;
q_2: consumed quantity of good 2;
γ: utility elasticity with respect to consumed good 1
$1-\gamma$: utility elasticity with respect to consumed good 2.

B) **Indirect utility.** Functions q_1 and q_2, which maximize utility, have the form:

$$q_1 = \gamma \frac{YM}{p_1}$$

$$q_2 = (1-\gamma) \frac{YM}{p_2}$$

where p_1 and p_2 are prices of goods and YM nominal income of consumers.

Replacing the above functions in the function of direct utility leads to the indirect utility function presented below:

$$UI = \frac{HYM}{p_1^{\gamma} p_2^{1-\gamma}}$$

with

$$H = \gamma^{\gamma} (1-\gamma)^{1-\gamma}$$

To facilitate the interpretation of indirect utility UI, let us recall that:

$$\frac{p_1^{\gamma} p_2^{1-\gamma}}{H}$$

is the composite price of a commodity basket when direct utility is equal to U.

In the case of perfect competition, elasticity γ corresponds to the fraction of bought commodity.

C) **Compensatory and equivalent variations.** If we assign an index "0" to the reference situation and index "1" to the new situation, then one may compare two welfare states in two manners:

a) $UI^0 = \dfrac{HYM^0}{(p^0{}_1)^\gamma (p^0{}_2)^{1-\gamma}}$

where UI^0 means the level of indirect utility at the reference situation.

b) $UI^1 = \dfrac{HYM^1}{(p^1{}_1)^\gamma (p^1{}_2)^{1-\gamma}}$

in the case of the new situation.

For further derivations, let us choose here, as a base of comparison, the new situation:

$$UI^0 - UI^1 = \frac{HYM^0}{(p^0{}_1)^\gamma (p^0{}_2)^{1-\gamma}} - \frac{HYM^1}{(p^1{}_1)^\gamma (p^1{}_2)^{1-\gamma}}$$

and, converting the above quantity into nominal values, we get compensatory variation CV:

$$\text{CV} = (UI^0 - UI^1)\frac{(p^1{}_1)^\gamma (p_2{}^1)^{1-\gamma}}{H} = \left(\frac{p^1{}_1}{p^0{}_1}\right)^\gamma \left(\frac{p^1{}_2}{p^0{}_2}\right)^{1-\gamma} (YM^0 - YM^1)$$

If CV < 0, then we have welfare improvement.

In simulation exercises through CGE models, the government first introduces a carbon tax to targeted polluting sectors. Simultaneously, it will reduce factor taxes in distorting production sectors. The next step is observing changes in price and on the household real income level. On this basis, one derives the compensatory variation CV presented above.

4 A Theoretical Example: CGE Model and Double-Dividend (DD)-Oriented Policies

In this section we will present a theoretical non-extensive cross-entropy (NCE) model to estimate parameters of an environmentally extended CGE model to assess impact of the DD hypothesis. The main issue to be underscored remains the rationale for applying the NCE approach in place of the traditional Kullback-Leibler cross-entropy model (Go et al., 2015). The response lies in the statistical properties of power law-related NCE. Due to the estimated parameters of constant elasticity of production models, outputs from both techniques must diverge with a higher performance in the case of NCE estimator (see V.5.3 for details). Both solutions will be similar only when the modelled phenomena display Gaussian distribution.

As presented earlier in this part of the book, the negative externalities resulting from pollution is one of the economy-distorting factors that prevents reaching a general equilibrium and Pareto optimum. Pearce (1991) proposed a model to generate double positive impact by introducing a tax imposition on pollutant activities. Once again, the first positive impact results in reducing pollutants through the increase of their production cost. If we use income generated by the tax imposition on pollutant goods to reduce tax in other sectors, a DD may result. This section presents a theoretical CGE model in the context of DD hypothesis testing. The proposed model can enable assessing to what extent the carbon tax can be identified as an important factor affecting the size of the DD, identifying the existence of a strong DD in the economy, or highlighting the weight of certain factors in affecting the presence and size of the DD.

This CE formulation may be written as follows:

$$Min \xrightarrow[p,r]{} \alpha_1 \frac{1}{q-1}\sum_k \sum_m p^E{}_{km}\left[(p^E{}_{km})^{q-1} - (q^E{}_{km})^{q-1}\right] -$$

$$\sum_k \sum_m (p^E{}_{km} - q^E{}_{km})(q^E{}_{km})^{q-1} +$$

$$\alpha_2 \frac{1}{q-1}\sum_t \sum_n \sum_j r_{tnj}{}^E\left[(r_{tnj}{}^E)^{q-1} - (s^E{}_{tnj})^{q-1}\right] - \sum_t \sum_n \sum_j (r_{tnj}{}^E - s^E{}_{tnj})(s^E{}_{tnj})^{q-1} \tag{5.11a}$$

Subject to:

$$F(X_t^E, Z_t^{0E}, Z_t^{uE}, B^E, \delta^E) = 0, \quad \forall t \in T$$

$$\delta^E = P(Z_t^E, B^E)$$

$$Y_t^E = G(X_t^E, Z_t^{0E}, Z_t^{uE}, B^E, \delta^E) + e_t \quad \forall t \in T$$

$$B_k^E = \sum_{m=1}^{M} p^E_{km} v_{km} \quad \forall k \in K$$

$$e^E_{tn} = \sum_{j=1}^{J} r_{tnj}^E w_{tnj} \quad \forall t \in T, n \in N$$

$$\sum_{m=1}^{M} p^E_{km} = 1, \quad \forall k \in K$$

$$\sum_{j=1}^{J} r_{tnj}^E = 1, \quad \forall t \in T, n \in N.$$

The above formulation (5.11a) is the same as the one presented in Part V (Equation 5.11), the only difference being the introduction in the model of an environmental sector. Consequently, this suggests the extension of the pre-existing model by adding environmental relations, as shown in the above optimization system where variables and parameters bear the superscript "*E*".

Depending on the target of the environmental model, the above CGE block F(..) = 0 will include additional endogenous variables (e.g., emitted CO_2, the dirty commodity, and its prices), exogenous variables like the environmental tax, the behavioral parameters like the elasticity of the environmental input, etc.

For example, to show the impact of the DD policy, one can add in the above CGE system F(..) = 0 the next equations, see (Sasmaz, 2016)[60] for counterfactual simulation purposes:

$$CO_{2t} = \alpha_{1t} + \beta_1 envtax_t + \beta_2 GDP_t + \beta_3 enecons_t + \zeta_{1t} \tag{6.1}$$

$$U\ nemp_t = \alpha_{2t} + \beta_4 envtax_t + \beta_5 GDP_t + \zeta_{2t} \tag{6.2}$$

Where CO_{2t} (Equation 6.1) is an endogenous variable which depends on the environmental tax rate $envtax_t$, the overall level of gross domestic production GDP_t, and the level of energy consumption $enecons_t$. Likewise, the endogenous variable U $nemp_t$ (Equation 6.2) explains the rate of unemployment which depends on the $envtax_t$ and the GDP_t. The index of time t is related to the targeted time series elements of the environmentally extended social accounting matrices Y_t^E. Parameters α_{1t} and α_{2t} are the constants. Parameters β_j (with j = 5) explain the long-run marginal change of the respective endogenous variable induced by a unit change of the explicative variable. Indeed, parameter interpretation in the model explains a long-run marginal change since we are dealing with an entropy model, the estimates of which will be generated by the maximum entropy principal rule under the CGE constraints. As noted in the previous section where the question was posed as to whether or not the maximum entropy solution is Pareto optimum, the obtained model solution may be different

60 Note that author has checked the DD hypothesis through an econometric panel data model.

from the solution of the traditional CGE optimal computation, which is not based upon a probabilistic distribution.

Besides the proposed illustrative model in (6.1–6.2), there exists a large formulation of the DD policy—oriented through CGE models (e.g., (Frasera & Waschik, 2013) (Takeda, 2007), (Taheripour, Khanna i Nelson, 2008), (Bento & Jacobsen, 2007)). As an interesting case study, Fraser and co-author presented in (Frasera & Waschik, 2013) a CGE model to assess a DD hypothesis where three taxes were examined: tax on the production of energy goods, on the production of carbon, and on the usage of carbon. To show the existence of DD for each tax revenue raised, authors proposed to offset pre-existing distortions in the CGE model through an endogenous tax that adjusts to keep constant government revenue. Finally, the author's outputs led to the existence of a strong DD associated with the existence of specific (immobile) factors in the production of energy goods. Note that these outputs were generated through a traditional CGE model.

5 Conclusions

In this section we have presented a theoretical NCE general formulation model in the case of general economic disequilibrium owing to, for example, environmental distortions. At the same time, the principal aspects of the empirical literature on the DD hypothesis has been presented. This represents a useful device for environmental economists who find in it a way to kill two birds with one stone, that is, reducing the carbon emission while creating better conditions for a balanced economic growth.

Future implementation of a PL-related NCE approach to estimate a CGE model, in general, and an environmentally extended CGE model, in particular, could reveal significant results. Indeed, the capacity of the PL-related NCE approach to handle non-linear inverse problem systems present in some sub-systems of the CGE model should produce positive outcomes.

Bibliography – Part VI

Adelman, I., & Robinson, S. (1978). *Income Distribution Policy in Developing Countries: A Case Study of Korea*. Stanford: Stanford University Press.

Armington, P.S. (1969). A theory of demand for products distinguished by place of production. *IMF Staff papers 16*, pp. 159–76.

Arrow, K.J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica, 22*, 265–90.

Arrow, K.J., & Hahn, F.H. (1971). *General Competitive Development Analysis* (40 ed.). San Francisco: Holden Day.

Arrow, K.J., Chenery, H.B., Minhas, B.S., & Solow, R.M. (1961). Capital-Labor Substitution and Economic Efficiency. *The Review of Economics and Statistics, 43*(3), pp. 225–250.

Bento, A.M., & Jacobsen, M. (2007). Ricardian rents, environmental policy and the 'double-dividend' hypothesis. *J. Environ. Econ. Manag., 53* (1), pp. 17–31.

Bottazzi G., G., Cefis, E., Dosi, G., & Secchi, A. (2007). Invariances and Diversities in the Patterns of Industrial Evolution: Some Evidence from Italian Manufacturing Industries. *Small Business Economics, 29*(1), pp. 137–159.

Bourguignon, F., Michel, G., & Miqueu, D. (1983). Short Run Rigidities and Long Run Adjustments in a Computable General Equilibrium Model of Income Distribution and Development. *Journal of Development Economics*, pp. 1–2.

Bwanakare, S. (2014). Non-Extensive Entropy Econometrics: New Statistical Features of Constant Elasticity of Substitution-Related Models. *Entropy*, pp. 2713–2728.

Bwanakare, S. (2016). Non-Extensive Entropy Econometrics and CES production Models: Country Case Study. *Journal of the International Association for Official Statistics (IAOS)*. Retrieved from http://content.iospress.com/articles/statistical-journal-of-the-iaos/sji1021

Capros, P., Karadeloglou, P., & Mentzas, G. (1990). An empirical assessment of macroeconometric and CGE approaches in policy modeling. *Journal of Policy Modelling, 12* (3), pp. 557–585. doi:http://dx.doi.org/10.1016/0161-8938(90)90013

Channing, A., Tarp, F., Robinson, S., Jensen, H.T., & Cruz, A. (2001). *Parameter estimation for a Computable general equilibrium model: A maximum entropy approach*. Washington: Working Paper, International Food Policy Research Institute.

Channing, A., Finn, T., Sherman, R., Jensen, H.T., & Cruz, A. (July1998). Social Accounting Matrices for Mozambique 1994 and 1995. Washington, D.C. 20006 U.S.A.: Working Paper, International Food Policy Research Institute.

De Melo, J. (1988). Computable General Equilibrium Models for Trade Policy analysis in Developing Countries: A Survey. *Journal of Policy Modeling, 4*, pp. 469–503.

Decaluwe, B., & Martens, A. (n.d.). Developing Countries and General Equilibrium Models: A Review of the Empirical Literature. *Journal of Policy Modeling, 10* (4).

Decaluwé, B., Martens, A., & Savard, L. (2001). *La politique économique du développement et les modèles d'équilibre général calculable.* Montréal, Canada: Les presses de l'université de Montréal.

Devarajan, S., & Robinson, S. (1983). *Selected Bibliography on Computable General Equilibrium Models of Developing Countries.* Harvard and U.C. Berkeley: Mimeo.

Donald, M., Scott, K., T., & Sherman, R. (2007). *Globe: A SAM Based Global CGE Model using GTAP Data*. Retrieved from www.usna.edu/EconDept/RePEc/usn/wp/usnawp14.pdf

Evans, G.W., & Ramey, G. (2006). Adaptive Expectations, Underparameterization and the Lucas Critique. *Journal of Monetary Economics, 53*, pp. 249–264.

Francois, J. (2001). Flexible Estimation and Inference Within General Equilibrium Systems. *Discussion Paper. 0129*. Adelaide 5005, Australia: Adelaide University.

Frasera, L., & Waschik, R. (2013). The Double Dividend hypothesis in a CGE model: Specific factors and the carbon base. *Energy Economics, 39*, pp. 283–295.

Gabaix, X. (2008). *Power laws in economics and finance*. Retrieved from http://www.nber.org/papers/w14299.

Giffin, A. (2009). From Physics to Economics: An Econometric Example Using Maximum Relative Entropy. *Physica A: Statistical Mechanics and its Applications,* pp. 1610–1620.

Ginsburg, V., & Waelbroeck, J. (1981). *Activity Analysis and General Equilibrium Modelling*. Amsterdam: North-Holland.

Ginsburgh, V., & Keyzer, M. (1997). *Structure of Applied General Equilibrium Models, 1997, MIT Press.* MIT Press.

Ginsburgh, V., & Waelbroeck, J. (1976). *Computational Experience with a Large General Equilibrium Model in computing equilibria.* Amsterdam North-Holland: How and Why (J. Los and M. Los).

Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2014). *Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach.* Retrieved from https://www.gtap.agecon.purdue.edu/resources/download/7136.pdf

Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015). Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach. *Development Economics, 52, part B,* pp. 790-811.

Golan, A., Karp, L.S., & Perloff, J.M. (1996). Esimating a mixed strategy employing maximum entropy. *Working paper, California Agricultural Experiment Station.* California.

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

Gossen, H.H. (1983). *The laws of human relations and the rules of human action derived therefrom*. Cambridge: MIT Press.

Green, W. (2003). *Basics of Econometrics.* NY: Hall Prencite, 5th edition.

Guerrien, B. (2000). *Dictionnaire d'analyse economique*. Paris, France: la Decouverte.

Harberger, A.C. (1959). The Corporation Income Tax: An Empirical Appraisal. *Tax Revision Compendium 1*, pp. 231–240.

Harberger, A.C. (1962). The Incidence of the Corporation Income Tax. *Journal of Political Economy 70*, pp. 215–240.

Harris, R., & Cox, D. (1983). Trade, Industrial Policy and Canadian manufacturing. *Canadian journal of economics*.

Herault, N. (2006). Building And Linking A Microsimulation Model To A Cge Model For South Africa. *South African Journal of Economics, 74*(1), pp. 34–58.
Ikeda, Y., & Souma, W. (2008). *International Comparison of Labour Productivity Distribution*. (Cornell University Library) Retrieved from arXiv:0812.0208v4
Jaynes, E.T. (1988). *Maximum- entropy and Bayesian methods in Science and Engineering, Vol. One: Fundations.* Boston, USA: G.J. Erickson and C.R. Smith, Kluwers Accademic Publishers.
Jevons, W.S. (1879). *The Theory of Political Economy.* London: Macmillan.
JIAN, X. (2000, August). An Environmentally Extended Social Accounting Matrix, Environmental and Resource Economics. *16*(4), pp. 391–406.
Jorgenson, D.W. (1984). *Econometric Methods for General Equilibrium Analysis In Applied General Equilibrium Analysis.* Cambridge: H. Scarf and J. Shoven, Eds., Cambridge University Press.
Jorgenson, D.W. (1998). *Growth. Volume 1: Econometric General Equilibrium Modeling.* Cambridge: The MIT Press.
Judge, G.G. & Mittelhammer, R.C. (2012). *An Information Theoretic Approach to Econometrics.* Cambridge, UK: Cambridge University Press.
Klump, R., & Papageorgiou, C. (2008). The CES Production Function in the Theory and Empirics of Economic Growth. *Journal of Macroeconomics, 2*, pp. 599–600.
Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*, 79–86.
Kydland, F.E., & Prescott, E.C. (1982). Time to Build and Aggregate Fluctuations. *Econometrica, 50* (6), pp. 1345–1370. doi:10.2307/1913386. JSTOR 1913386
Kydland, F., & Prescott, E.C. (1977). Rules Rather than Discretion: The Inconsistency of Optimal Plans. *Journal of Political Economy*, pp. 473–492. doi:10.1086/260580
Lee, D., & Misiolek, W. (1986). Substituting Pollution Taxation for General Taxation: Some Implications for Efficiency in Pollution Taxation. *JEEM 13,* pp. 338–47.
Leontief, W. (1941). *The Structure of American Economy, 1919–1929.* Cambridge: Harvard University Press.
Lofgren, H., Harris, R.L., & Robinson, S. (2000). *A Standard Computable General Equilibrium (CGE) Model in GAMS.* Washington: Intl Food Policy Res Inst.
Lucas, R. (1976). Econometric Policy Evaluation: A Critique. *1*, 19–46. doi:10.1016/S0167–2231(76)80003–6
Mansur, A., & Whalley, J. (1984). Numerical Specification of Applied General Equilibrium Models: Estimation, Calibration, and Data in Scarf and Shoven. *Applying General Equilibrium Analysis*, pp. 69–127.
Mantegna, R.N., & Stanley, H.E. (1999). *Introduction to Econophysics: Correlations and Complexity in Finance.* Cambridge: Cambridge University Press.
Marshall, A. (1890). *Principles of Economics.* London: Macmillan.
McKenzie, L.W. (1959). On the Existence of General Equilibrium for a Competitive Economy. *Econometrica, 1*, pp. 54–71. doi:10.2307/1907777. JSTOR 1907777
McKenzie, L.W. (1981). The Classical Theorem on Existence of Competitive Equilibrium. *Econometrica, 4*, pp. 819–841. doi:10.2307/1912505. JSTOR 1912505
McKenzie, L.W. (1987). Turnpike theory. *The New Palgrave: A Dictionary of Economics*, pp. 712–720.
McKenzie, L.W. (1954). On Equilibrium in Graham's Model of World Trade and Other Competitive Systems. *Econometrica, 2*, pp. 147–161.
Muth, J.F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica, 3*, pp. 315–335. doi:10.2307/1909635
Nash, J.F. (1950). The Bargaining Problem. *Econometrica, 2*, pp. 155–162. doi:10.2307/1907266. JSTOR 1907266
Pearce, D. (1991). The Role of Carbon Taxes in Adjusting to Global Warming. *The Economic Journal*(101), pp. 938–948.

Pigou, A.C. (1932). *Economics of Welfare*. London: Macmillan Publishing, 4th edition.
Pukelsheim, F. (1994). The Three Sigma Rule. *American Statistical Association, 2*, pp. 88–91.
Pyatt, G., & Thorbecke, E. (1976). *Planning Techniques for a Better Future,*. Geneva, Switzerland: International Labour Office.
Pyatt, G. (1988). A SAM approach to modeling. *Elsevier, Journal of Policy Modeling, 10*(3), pp. 327–352.
Pyatt, G., & Roe, A. (1977). *Social Accounting Matrices for Development Planning: With Special Reference to Sri Lanka*. Cambridge: Cambridge University Press.
Sasmaz, M.U. (2016). Validity of double dividend hypothesis in EU–15 countries. *Global Journal on Humanites & Social Sciences*, pp. 30–36.
Scarf, H. (1984). *The Computation of Equilibrium Prices*. Cambridge: (H. Scarf and J. Shoven, Eds.), Cambridge University Press.
Scarf, H. (1969). An Example of an Algorithm for Calculating Equilibrium Prices. *American Economic Review, 59*.
Shoven, J.B., & Whalley, J. (1972). A General Equilibrium Calculation of the Effects of Differential Taxation of Income from Capital in the U.S. *Journal of Public Economics 1 (3–4)*, pp. 281–321.
Shoven, J.B., & Whalley, J. (1992). *Applying General Equilibrium*. New York: Cambridge University Press.
Sims, C.A. (1987). A Rational Expectations Framework for Short-Run Policy Analysis. *New Approaches to Monetary Economics*, pp. 293–308.
Sims, C.A. (2007). Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian. Princeton : Princeton University.
Stone, R. (1955). *Input-Output and the Social Accounts*. New York: J. Wiley.
Stone, R., & Brown, A. (1962). *A computable model for economic growth*. Cambridge, UK: Cambridge Growth Project.
Taheripour, F., Khanna, M., & Nelson, C.H. (2008). Welfare impacts of alternative public policies for agricultural pollution control in an open economy: a general equilibrium framework. *Am. J. Agric. Econ., 90*(3), pp. 701–718.
Takeda, S. (2007). The double dividend from carbon regulations in Japan. *J. Jpn. Int. Econ., 21*, pp. 336–364.
Terkla, D. (1984). The Efficiency Value of Effluent Tax Revenues. *Journal of Environmental Economics and Management, 11*(2), pp. 107–123.
Tovar, C. (2009). DSGE Models and Central Banks. *Economics, 3 (2009–16)*(1). doi:10.5018/economics-ejournal.ja.2009–16
Tsallis, C. (2009). *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*. Berlin: Springer.
Tullock, G. (1967). Excess Benefit. *WaterR Resources Research, 3*(2), pp. 643–644.
Walras, L. (1874). *Éléments d'économie politique pure, ou, Théorie de la richesse sociale*. Lausanne: L. Corbaz & Cie.

Part VII: **Concluding Remarks**

The purpose of this book is to present a non-extensive entropy econometrics technique as a new, robust devise of economic ill-behaved inverse problem modelling when traditional econometrical approaches may not work. Particular attention has been paid to national accounts tables supposed to display an endogenous general equilibrium of the economy. Non-extensive entropy related power law has proven to ubiquitously describe a large variety of phenomena, including those of financial markets. This book has tried to show that economic phenomena should also be described by this same law which generalises Shannon entropy and thus displays more stable solutions, including those beyond the Gaussian framework.

Nevertheless, for a new approach to disserve attention, it should be consistent with an existing theory and be proven to lead to new empirical advantages. Consequently, the second part of this study has been devoted to the main theoretical techniques for solving inverse problems. In particular, starting with the Kullback-Leibler statistical information definition, we have presented the principal inverse problem solution techniques—and their limits—before characterizing Tsallis entropy with respect to its higher generalizing qualities. Next, based on recent Shannon entropy econometrics techniques, we have proposed a non-extensive entropy econometrics approach which encompasses the Gibbs-Shannon entropy case. In terms of their probabilistic definition, both econometrical techniques are characterized by corresponding functional forms in criterion function. In both cases, constraining parts are based on the Bayesian method of moments. Nevertheless, due to the non-extensivity of Tsallis entropy, that aspect is magnified through a particular treatment of constraining moment equations, such as the use of escort distribution. Up to now, those forms of restriction have been used in statistical thermodynamics, so their successful transfer to entropy econometrics—then to the social sciences in general —constitutes the first intermediary result of this book. This is sufficiently important, seeing that, on statistical thermodynamics grounds, the discussions concerning the applicability conditions of different forms of restrictions are not yet closed. On the other hand, one cannot exclude that their application in social science could help in better understanding their use in statistical thermodynamics. In spite of sparse literature, the second intermediary result remains the formal connection of Shannon-Gibbs based Kullback-Leibler statistical theory of information to the proposed Tsallis non-extensive entropy econometrics. As mentioned in later sections of the book, such a link represents a key element for empirical investigations since it will enable measuring divergence between hypotheses when non-ergodic phenomena are analysed.

Until now, econophysicists have applied non-extensive entropy to high frequency phenomena. This book, through a probabilistic and dynamic characterization of power law—using scaling processes—puts forward the possibility of non-extensive entropy econometric modelling with low-frequency data, for instance, annual statistical series. This is significant since we then open the possibility of modelling with data reflecting real world economic practices. A central achievement of this work remains the proposed and demonstrated theorem linking economics to statistical thermodynamics through the inheritance properties of PL. We then endeavoured to build a stronger bridge between these two disciplines.

When the non-extensive entropy procedure is used to update and forecast input-output tables, outputs remain similar to those obtained using traditional approaches, such as Shannon entropy or the RAS method. If our purpose had been to prove that the proposed new Tsallis entropy-based approach worked, obtaining similar outputs would have been sufficient. However, our target was more ambitious. We needed to show the superiority of the new approach. At least two reasons are behind this limited achievement. The first is that the structure of input-output-based tables may reflect Gaussian law, i.e., the convergence limit of power law. Consequently, Shannon and Tsallis entropies should display the same outputs. This is what has occurred in our study. As far as the RAS approach is concerned, achieving a higher performance by entropy techniques would have required, according to the existing literature, a higher number of consistent restrictions or stochastic modelling conditions. We have presented an example with powerful results when we were forecasting an aggregated input-output table of 27 EU countries. This interesting but time-consuming research disserves further investigation to better set modalities and conditions of each of the above rival approaches. In the context of these input-output table-based models, the next principal finding of this work remains the new proposed non-extensive entropy econometrics technique for estimating a large class of nonlinear stochastic models considered non-tractable. We have particularly in mind a large class of non-stationary or long memory econometric models, the data of which, at first glance, could correspond with an unknown time scaling intermediary level of power law.

The above findings were necessary to show the possibilities of extending generalized non-extensive Tsallis entropy to low frequency econometric models and, in general, to expand the estimation possibilities of econometric modelling, including static or dynamic CGE models and their extensions. Since data aggregating level—as evidenced by recent literature—may correspond to a certain probabilistic family of law, this fact could represent a paradigm shift in many fields, including applied statistics and econometrics.

Appendix

Annex C. Computational Aspects of Using GAMS

1.1 Introduction to GAMS

During this presentation, we have solved most of problems using the GAMS (General Algebraic Modelling System) code available on the site www.gams.com. A free student version is available on that site. Different commercial copies for different system platforms are available, too. In early 2003, a new *GAMS User Guide* was released, expanding on the earlier Guides by Brooke, Kendrick, and Meeraus and, later, Ramen. Succinctly speaking, GAMS is a language for setting up and solving mathematical programming optimization models. It is a compact language simultaneously allowing one to specify the structure of an optimization model, specify and calculate data for the model, solve the model, conduct report writing on a model and perform a comparative static analysis. Any introductory GAMS user guide provides the different steps in programming in that language.

Steps of programming in GAMS:

1. Variable specifications
2. Equation specifications
   a. declaration
   b. algebraic structure specification
3. Model statement
4. Solve statement

To be more complete, an example of GAMS programming, presented by Dhazn Gillig & Bruce A. McCarl (Department of Agricultural Economics, Texas A&M University) at http://agecon.tamu.edu/faculty/mccarl/mccarl.htm is included below.

1.2 Formulation of a Simple Linear Problem in GAMS

a) Mathematical model:

Maximize	$109X_1$	$90X_2$	$115X_3$	
Subject to	X_1	X_2	X_3	≤ 100
	$6X_1$	$4X_2$	$8X_3$	≤ 500
	X_1	X_2	X_3	≥ 0 (non-negative)

b) Structure of the program:

```
VARIABLES
    Z              Variable Z         ;

POSITIVE VARIABLES
    X1             Variable X1
    X2             Variable X2
    X3             Variable X3        ;

EQUATIONS
    Equation1              Equation 1
    Equation2              Equation 2
    Equation3              Equation 3        ;

Equation1..
    Z =E= 109*X1 + 90*X2 + 115*X3            ;

Equation2
    X1 + X2+ X3 =L= 100                      ;

Equation3
    6*X1 + 4*X2 + 8*X3 =L= 500               ;

MODEL Example1 /ALL/;

SOLVE Example1 USING LP MAXIMIZING Z;
```

As is easy to observe, the sequences of solving the above linear program problem are ordered according to the above GAMS programming steps.

1.3 Application to Maximum Entropy Models

1.3.1 The Jaynes Unbalanced Dice Problem

This problem has been presented in Part II of this book and outputs displayed in Table 2.1. Here we present a code in GAMS (see Golan et al.,1996) for solving such a simple unidimentional problem.

**Derivation of probability structure of unfair dice with the GAMS code*:

```
 SET
     i index/1*6/
```

```
parameter
    x(i) support/1 1
         2 2
         3 3
         4 4
         5 5
         6 6/;
POSITIVE VARIABLE P(i) probabilities;
VARIABLE OBJ objective;
EQUATIONS
    OBJECTIVE  entropy objective
    ADD        additivity constraint
    CONSIST    consistency equation;
OBJECTIVE..OBJ =E= -sum(i, P(i)*log(1.e-9+P(i)));
ADD..SUM(i, P(i)) =E= 1;
CONSIST..SUM(i,(x(i)*P(i))) =E= 4.5;
Model classoc/ALL/;
Solve classoc maximizing OBJ using NLP;
DISPLAY P.L;
DISPLAY OBJ.L;
```

1.3.2 Non-Extensive Entropy Econometric Model

This is a more complex example written by S. Bwanakare (2009) for real world problems, in this case a labour demand model for the Polish economy. More explanations are provided in the following introductory text of the program

$ontext

Generalized Maximum Entropy Parameter estimation of a labour demand econometric model for Podkarpacki province. Moving along rationale expectation mainstreams, this model estimates long-run and short-run impact of demand labour determinants. However, since this model belongs to the class of ARDL (autoregressive distributed lag) model, exogenous variables are not independent of the error term. Additionally, due to the small data sample available for the Podkarpacki district, assumptions concerning random term distribution becomes unknown. Then, estimation of parameters using classical methods (like the LS) become ineffective. The model below exploits the Generalized Maximum Entropy principle to estimate parameter of the labour demand model.

$offtext

```
SET
   M index /1*5/
   J index /1*5/
   t index/1*9/
   k index /lnLt,lnYt,lnLt_1,t,ao/
   xk(k) index /lnYt,lnLt_1,t,ao/
;
table data(k,T)

          1       2       3       4        5        6        7        8       9
LnLt      0.0049  0.0062  0.0026  -0.0163  -0.0373  -0.0137  -0.0112  0.0001  0.0051
LnYt      0.0936  0.0875  0.0626  0.0301   0.0343   0.0245   0.014    0.0212  0.0343
LnLt_1    4.5126  4.6013  4.6826  4.7426   4.7890   4.8606   4.8988   4.9240  4.9451
t         1       2       3       4        5        6        7        8       9
ao        1       1       1       1        1        1        1        1       1

display data;
*$offtext ;
Positive variables
        P(xk,M) parameter probabilities
        W(T,j) error probabilities
        sigm standard error on parameters
        q  parameter tsallis   ;
parameter
       V(t,j) support space for error
       Par(xk) parameter estimates
*       sigm standard error on parameters
       X(T,xk) explanatory variables
       Y(T,*) dependent variables
       XY(T,k) all variables
       kurt(k) forth moment


*       epsilon positive small real
       sigmaa(xk) standard error on all variables
        sigmaa(xk) sample standard error
                    /lnYt    0.0
                     lnLt_1  0.0
                       t     0.050 /
        Z(M) parameter support /1 -1.000
                                  2 -0.500
                          3   0
```

```
                    4   0.500
                  5  1.000/;

               v(t,"1")  =-3*sigmaa("lnYt") ;
               v(t,"2")  = -1*sigmaa("lnYt");
               v(t,"3") =  0*sigmaa("lnYt");
               v(t,"4") =  1*sigmaa("lnYt");
               v(t,"5") =  3*sigmaa("lnYt") ;
               v(t,"1") =-3*sigmaa("lnLt_1");
               v(t,"2") = -1*sigmaa("lnLt_1");
               v(t,"3") =0*sigmaa("lnLt_1");
               v(t,"4") = 1*sigmaa("lnLt_1");
               v(t,"5") = 3*sigmaa("lnLt_1") ;
               v(t,"1") =-3*sigmaa("t");
               v(t,"2") = -1*sigmaa("t");
               v(t,"3") =0*sigmaa("t");
               v(t,"4") = 1*sigmaa("t");
               v(t,"5") = 3*sigmaa("t") ;

Display v;
        W.l(t,"1")=1/72 ;
        W.l(t,"2")=27/72 ;
        W.l(t,"3")=16/72 ;
        W.l(t,"4")=27/72 ;
        W.l(t,"5")=1/72;
 parameter
       epsilon /0.0001/ ;
        W.lo(t,j)=epsilon ;
        W.up(t,j)=1 ;
        p.lo(xk,m)=epsilon;
        p.UP(xk,m)=1 ;
*        q.lo=epsilon;
Variable OBJ objective ;
Y(t,"lnLt") = data("lnLt",t);
X(t,"lnYt") =data("lnYt",t);
X(t,"lnLt_1") =data("lnLt_1",t);
X(t,"t") =data("t",t);
X(t,"ao") =data("ao",t);
q.l=1.5;
display Y, X;
Equations
       OBJECTIVE objective function
```

```
        ADD1(xk)  parameter additivity constraints
        ADD2(T)  error additivity constraints
        CON(T)   consistency constraints;
OBJECTIVE..OBJ  =E=  sum(xk,  sum(M,  (P(xk,M)-P(xk,M)**q)/(q-1))+sum(T,Sum(J,
(W(T,j)-(W(T,j)**q)/(q-1)))));
ADD1(xk)..sum(m, P(xk,M)) =E=1;
ADD2(T)..sum(J, W(T,J)) =E=1;
CON(T)..sum(xk,   X(T,xk)*sum(M,   [(P(xk,M)**q)/sum(M,   (P(xk,M)**q)]*Z(M)))
+sum((J,[W(T,j)**q)/sum((J,W(T,j)**q)]*V(t,j)) =E=Y(T,"lnLt");
=E=Y(T,"lnLt");
Model labour /ALL/;
labour.optfile = 1 ;
labour.HOLDFIXED = 1 ;
labour.scaleopt=1 ;
option NLP = minos5;
Solve labour maximising OBJ using NLP;
PAR(xk) = sum(M, P.L(xk,M)*Z(M));
DISPLAY PAR,q.l;
*in sample teoretical yy
parameter

logLL(T) in sample pronostical Y ;
logLL(T) = sum(xk, X(T,xk)*Par(xk)) ;
display logLL;
parameter
god goodness coefficient
S(xk) parameter information index
et(T) estimation of random term;

S(xk)=(1-sum(m,P.L(xk,M)**q.l))/(q.l-1)*(1-q.l)/[5**(1-q.l)-1];
et(T)= data("lnLt",T)-logLL(T);
god=obj.l*(1-q.l)/(11*[5**(1-q.l)-1]);

display obj.l,god, et, S, p.l,w.l;
```

Annex D. Recovery of Pollutant Emissions by Industrial Sector and Region: an Instructional Case

1.1 Introduction

In this section, we are going to present a case that, unlike the case studies carried out in previous parts of the monograph, shows the limits of the entropy econometrics approach when some initial conditions are not fulfilled while dealing with an inverse problem. Thus, the example illustrates what should not be done to reach reliable model estimates. This is important for the less experienced modellers who try to recover information on the basis of this theory. In this instance, we see that an important property of respective continuity of probability measures of two hypotheses, as those were defined in part II of the book, is missing. In such a case, more statistical data should be collected to allow for an inference based on regular conditions. In this example, we use a GAMS code.

1.2 Recovery of Pollutant Emission by Industrial Sector and Region: the Role of the *Prior*[61]

Suppose we dispose of information on aggregated greenhouse emissions simultaneously at the regional and sectorial level. In Table 23 below, we can depict this total information in the last column and last row, respectively. It is important to underscore here that these two pieces of information are from two different sources and, plausibly, were initially collected for different purposes. Thus, we have to join these two pieces with the hope of recovering the real distribution state of pollutant emission by industrial sector and region. In this example, there is an additional complication related to the impossibility of comparing these two totals. In fact, values in the last column constitute an unknown aggregate of standardized quantities of different pollutants with different units of measure. Inversely, in the last row, available statistical data of aggregates have been converted into tons of CO_2 equivalent. Suppose, in this problem, we are asked to recover the regional and sectorial level of greenhouse emissions in tons of CO_2 on the basis of the above availability of statistical data.

If column and row totals were explained in the same unit of measure, we would then have to deal with a direct extension of the Jaynes dice problem (see Table 1 in Part II) from one to two-dimensional discrete space. In this context, finding a unique optimal solution for such a problem requires finding an equation system with a

61 For a complete solution of this problem, we send the reader to: *Acta Physica Polonica A* 2015, 127, A-13. [CrossRef] closed form. Additional and consistent *a priori* information is required to enable an optimization entropy device moving from uniform distribution toward an optimal, global solution. For a researcher unacquainted with solving this kind of problem, the present example aims at revealing the limitations of the maximum entropy principle when minimum conditions are not fulfilled. In this example, the *a priori* is uniformly distributed information of pollutant emissions by industrial sector and region. Doing so, we try then to solve the problem using the well-known principle of insufficient reason already discussed in the introduction of this work.

As already observed, in the present problem, we have in column and row totals different units of measure. In this case, we suggest introducing a scaling factor (as an additional new *a priori*) in consistency equations of the model, theoretically allowing the table to then balance.

The Model

According to what has been said above and then without *a priori* information about the system, we propose applying non-extensive maximum entropy in the criterion function under traditional restrictions such as moments and normality conditions. The model takes the following usual form:

$$\max\left(H_q(p)\right) = \left[1 - \sum_i \sum_j \left(p_{ij}\right)^q\right] \Big/ (q-1)$$

s.t.:

$$Y_i = c(i) \sum_j X_j p_{ij}$$

$$x_{ij} = X_j \frac{p_{ij}^{\ q}}{\sum_i p_{ij}^{\ q}}$$

$$1_j = \sum_i \frac{p_{ij}^{\ q}}{\sum_i p_{ij}^{\ q}}$$

where:
Y_i: means total by row i,
X_j: means total by column j,
p_{ij}: probabilistic structure of greenhouse emissions by region
x_{ij}: quantity of pollutant emission inside cells.

In this problem, we dispose of up to 122 points to be estimated (then including an additional 17 scaling factors $c(i)$) and only 23 observation points. Naturally, we are dealing with an inverse problem that we know how to solve from theoretical and

Table 23: Total pollutants emission by industrial sector and by region (before estimation)

Regions\Energy industry	construction	Transport	Industrial processing	Agriculture	Waste	Other	Total GreenhouseGas (standardized)	
Lodz								3.88928103
Mazowieckie								12.2540834
Małopolskie								-0.0723657
Śląskie								8.8672366
Lubelskie								-0.82089026
Podkarpackie								-3.41721603
Podlaskie								-3.92242795
Świętokrzyskie								-4.41226141
Lubuskie								-5.9413236
Wielkopolskie								5.7919945
Zachodniopomorskie								-3.02668716
Dolnośląskie								0.67034141
Opolskie								-4.61146968
Kujawsko-Pomorskie								-1.11131699
Pomorskie								-1.83960985
Warmińsko-Mazurskie								-2.29736824
Total tons CO_2 equivalent	180369	32469	36443	28877	34581	9437	67787	-16.0709987

Source: own calculation on the basis of EUROSTAT information

Table 24: Total pollutant emission by industrial sector and region (post entropy estimation).

	indusener	indutrans	trans	procind	agr	dech	aut	Total
lod	445.943	0.00E+00	0.00E+00	0.00E+00	92.881	0.00E+00	0.00E+00	538.824
maz	25838.76	0.00E+00	15494.79	3373.983	2457.911	1027.966	0.00E+00	48193.4
mal	10463.45	1201.047	760.303	0.00E+00	993.605	4634.521	10796	28848.92
sla	7265.899	1201.05	10.309	0.00E+00	0.00E+00	0.00E+00	44955.66	53432.92
lub	26102.65	1201.046	2023.227	0.00E+00	0.00E+00	0.00E+00	0.00E+00	29326.92
pod	11110.19	1201.047	1.773	0.00E+00	2289.638	0.00E+00	0.00E+00	14602.65
podl	7888.465	2031.926	20.356	0.00E+00	753.312	0.00E+00	942.165	11636.22
swi	3001.472	1171.582	1934.917	0.00E+00	2240.56	0.00E+00	298.363	8646.894
lubu	372.657	3300.148	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3672.805
wiel	2985.15	2914.626	0.00E+00	23710.43	15831.89	2197.649	7562.365	55202.12
zac	9952.093	691.237	1130.836	1792.583	2400.436	1352.623	250.21	17570.02
dol	21697.25	1443.422	9855.435	0.00E+00	2401.033	224.241	995.144	36616.52
opo	7448.701	827.828	3820.23	0.00E+00	317.823	0.00E+00	0.00E+00	12414.58
kuj	25001.55	0.00E+00	1.17	0.00E+00	2401.077	0.00E+00	993.112	28396.91
pom	20794.83	0.00E+00	5.611	0.00E+00	2400.832	0.00E+00	993.983	24195.26
war	0.00E+00	15284.04	1384.048	0.00E+00	0.00E+00	0.00E+00	0.00E+00	16668.09
Total	180369	32469	36443	28877	34581	9437	67787	

Source: own calculation.

empirical models encountered earlier in this work. Outputs from derivations are shown in Table 24. Only a limited part of the information has been recovered. In fact, column totals have been accurately recovered, and line totals seem, to some extent, to conform to our expectations. In fact, we see that predominant regions in greenhouse gas emissions, like Mazowieckie, Slaskie, seem to produce the highest levels of pollutants.

As already known, adding a new piece of *a priori* information far from uniform distribution and consistent with properties of the system would significantly improve the quality of prediction of these emission quantities. Additionally, as has been observed in the case of forecasting input-output systems (Part III), entropy estimators belong to the family of Stein estimators. As such, smaller probabilities are shrunk and higher probabilities then dominate in the space of solutions. Fortunately, adding more *a priori* information in the model enhances parameter precision and then allows for recovering the influence of smaller events.

Index of Subject

a carbon tax xiv
a computable general equilibrium (CGE) model xiv
a prior support 5
a social accounting matrix xiii
balancing a social accounting matrix xiv
auto-organized systems 2
Bayesian efficient processing rule 5
Bayesian rule 16
constant elasticity of substitution 155, 157, 158, 159, 162
Cramer-Rao inequality 3
divergence between hypotheses 19
entropic Tsallis-q complexity parameter 8
Gaussian basin of attraction 8, 35
generalized method of moments 10
Gibbs-Shannon maximum entropy 2, 25
high frequency series 4
ill-behaved inverse problem 10
ill-behaved matrix 10
ill-behaved matrix 88, 92
ill-behaved matrix 3
Information-theoretic 2
input- output table 11
inverse problem 2, 3, 10, 11, 22, 23, 25, 27, 30, 38, 42, 54, 62, 75, 78, 88, 89, 94, 101, 106, 204, 205
Kullback-Leibler cross-entropy
cross-entropy 98
Kullback-Leibler information divergence xiii
Levy instable law 158
Levy instable phenomena 7
low frequency series 6, 8
maximum entropy 2, 4, 5, 6, 10, 11, 20, 22, 23, 27, 28, 39, 40, 41, 54, 56, 75, 90, 91, 103, 122, 127, 150, 151, 152, 199, 205
Maximum entropy principle 2
maximum likelihood 10, 162
maximum likelihood method 10
Moore-Penrose generalized inverse xiii
Moore Penrose pseudo inverse 23
multi fractal properties 4
multinomial 25
multinomial distribution 25
national account 10, 11, 23, 42, 54, 55, 75
non-additive statistics 22
non-ergodic systems 101, 158, 160
non-extensive entropy 22, 30, 63, 169
non- representative sample 2
optimum Pareto 11, 140, 142
posterior 17, 18, 75, 101, 105, 106
posterior probability 17
power law 27, 29, 30, 62, 78, 100, 105, 122, 123, 124, 126, 127, 128, 129, 157, 158
principle of sufficient reason 2
prior 3, 5, 17, 29, 40, 75, 98, 99, 100, 101, 104, 106, 122, 152, 156, 157, 163
prior probability 17
q-Tsallis 31
scale-invariant
scale-invariant structures 29
scale-invariant structures 29
social accounting matrix 11, 75, 88, 92, 93, 95, 96, 101, 125, 134, 143, 177
statistical theory of information 54, 75, 160
support space 5, 10, 16, 18, 19, 39, 40, 41, 63, 99, 103, 104, 107, 161, 163, 201
Tsallis non-additive statistics xii

Index of Authors

Bayes 2, 17, 105, 106
Bernoulli 2
Boltzmann 2, 41
Borges 168
Bottazzi 105
Bregman 2, 36, 37, 38, 101, 157
Brooke 198
Bruce 176, 198
Buchanan 176
Cauchy 2
Champernowne 105
Channing 151
Cox 2
Curado 37
Downs 176
Eurostat 73
Foley 91
Fujiwara 7
Gibbs 2, 4, 7, 8, 16, 21, 22, 25, 26, 27, 30, 31, 33, 62, 76, 94, 152, 156
Ginsburgh 138
Golan 3, 4, 6, 28, 29, 38, 99, 102, 152, 158
Gosh 60
Gossen 137
Granger 125
Harberger 135
Harris 135
Hartley 16, 18
Heckelei 4
Ikeda 7
Jaynes 2, 3, 5, 16, 21, 22, 25, 26, 27, 75, 100, 102, 199, 204
Jeffreys 2
Jevons 137
Judge 3, 29
Kantelhardt 4
Klump 159
Kullback 2, 3, 8, 10, 16, 17, 22, 27, 28, 29, 35, 36, 37, 49, 50, 63, 88, 98, 99, 100, 101, 157, 158
Kullback-Leibler 4, 8, 10, 22, 27, 36, 37, 49, 50, 88, 98, 99, 100, 101, 157, 158
Laplace 2
Leibler 3, 27, 35, 36, 37, 63, 75, 100
Leontief 57, 58, 134, 150, 163
Lévy 2
Lindley 3
Lucas 134
Mandelbrot 2
Mansur 94
Mantegna 105
Maxwell 2
Mc Donald 144
Menger 137
Miller 3, 29
Misiolek 184
Mounfield 7
Nordhaus 176
OECD 93
Ormerod 7
Papageorgiou 159
Papoulis 26
Pearce 184
Pigou 183
PL 3, 4, 6, 7, 8, 11
Pukelsheim 107, 163
Pyatt 105, 143
Renyi 2, 31, 33, 35
Robinson 94, 143, 152
Scarf 135
Scott 146
Shannon 2, 3, 4, 5, 7, 8, 10, 16, 21, 22, 25, 26, 27, 29, 30, 31, 33, 35, 41, 62, 75, 76, 94, 97, 98, 152, 165
Shoven 151
Souma 7
Stanley 7, 105, 126
Stone 55, 135
Terkla 184
Theil 24, 97
Thierfelder 146
Tsallis 2, 6, 7, 8, 16, 21, 22, 27, 29, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 50, 54, 76, 78, 88, 100, 101, 103, 104, 106, 157, 158, 160, 163, 164, 165, 169
Tullock 183
United Nations 55
Weierstrass 2
Whalley 94

Corrections to second edition

Page	1st edition	corrected edition
vii	Tsallis Entropy and mainproperties	Tsallis Entropy and main properties
7	Shlesinger (Shlesinger, Zaslavsky, & Klafter, Strange Kinetics, 1993)	Shlesinger, Zaslavsky, & Klafter, Strange Kinetics, 1993.
7	(Shlesinger & et al, Lévy Flights and Related Topics in Physics, 1995).	Shlesinger et al., Lévy Flights and Related Topics in Physics, 1995.
17	variableis	variable is
19	$I(\mu_1:\mu_2)=f_2(x)\log\frac{f_2(x)}{f_1(x)}d\lambda(x)$ (2.8) as the mean information from μ_2 for discrimination in favour of H_2 against H_1, one can define the divergence (noted ∇) by: $\nabla(H_1, H_2) = I(\mu_1 : \mu_2) + I(\mu_2 : \mu_1) =$ $\int (f1(x)-f_2(x))\log\frac{f_1(x)}{f_2(x)}d\lambda(x) =$ $=\log\frac{P(H_1\mid x)}{P(H_2\mid x)}d\mu_1(x)-\int\log\frac{P(H_1\mid x)}{P(H_2\mid x)}d\mu_2(x)$ divergence between hypotheses. (2.9)	$I(\mu_1:\mu_2)=f_2(x)\log\frac{f_2(x)}{f_1(x)}d\lambda(x)$ (2.8) as the mean information from μ_2 for discrimination in favour of H_2 against H_1, one can define the divergence between hypotheses (noted ∇) by: $\nabla(H_1, H_2) = I(\mu_1 : \mu_2) + I(\mu_2 : \mu_1) =$ $\int (f1(x)-f_2(x))\log\frac{f_1(x)}{f_2(x)}d\lambda(x) =$ $=\log\frac{P(H_1\mid x)}{P(H_2\mid x)}d\mu_1(x)-\int\log\frac{P(H_1\mid x)}{P(H_2\mid x)}d\mu_2(x)$ (2.9)
26	trialswe	trials we
29	Tsallis entropy and mainproperties	Tsallis entropy and main properties
33	α	λ
35	Kullback-Leiber	Kullback-Leibler
40	PM	P_m
51	Golan, A., Judge, G.G., & Perloff, M.J. (1996c)	Golan, A., Judge, G.G., & Perloff, J.M. (1996c)
56	Pj	P_j
63	Min(p,p0)	Min(p, p_0)
74	*IOtable^p*	I-Otable^p
75	*IOtable^p*	I-Otable^p
78	Min(b1,b0)	Min(b_1, b_0)
83	Almon, & Clopper. (2000).	Almon, C. (2000).
83	Avonds, & Luc. (2007).	Avonds, L. (2007).

Page	1st edition	corrected edition
108	"... those of of Salem..."	"... those of Salem...."
117	*ceteris per ibis*	*ceteris paribus*
130	Miller Ana Carina, Alan Matthews, Trevor Donnellan, A. (2005).	Miller, A.C., Matthews, A., Donnellan, T., & O'Donoghue, C. (2005).
138	Mc Kenzie	McKenzie
138	Cobb Douglas	Cobb-Douglas
143	(Pyatt and Graham, 1988)	(Pyatt, 1988)
144	Cobb Douglas	Cobb-Douglas
153	Darwin Fowler	Darwin-Fowler
155	Go et al., 2014	Go et al., 2015
170	Donald, M., Scott, K., T., & Sherman, R. (2007).	McDonald, S., Thierfelder, K., Robinson, S. (2007).
171	Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015). Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach,. Development Economics, Prospects Group, Waszyngton DC. Waszyngton: World Bank Group, Development Economics, Prospects Group.	Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015). Estimating parameters and structural change in CGE models using a Bayesian cross-entropy estimation approach. Policy Research working paper no. WPS 7174.
171	A Standard Computable General Equilibrium (CGE) Model in GAMS. Waszyngton:	A Standard Computable General Equilibrium (CGE) Model in GAMS. Washington:
172	McKenzie, & Lionel , W. (1959).	McKenzie, L.W. (1959).
172	McKenzie, & Lionel , W. (1981).	McKenzie, L.W. (1981).
172	McKenzie, & Lionel , W. (1987).	McKenzie, L.W. (1987).
172	McKenzie, L. (1954).	McKenzie, L.W. (1954).
172	Pyatt, & Graham. (1988).	Pyatt, G. (1988).
181	TAXPOLL	TAX_{POLL}
190	Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015, January 2015, Waszyngton DC). Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach,. Development Economics, Prospects Group, January 2015, Waszyngton DC. Waszyngton: World Bank Group, Development Economics, Prospects Group.	Go, D.S., Lofgren, H., Ramos, F.M., & Robinson, S. (2015). Estimating Parameters and Structural Change in CGE Models Using a Bayesian Cross-Entropy Estimation Approach. Development Economics, 52, part B, pp. 790-811.
191	A Standard Computable General Equilibrium (CGE) Model in GAMS. Waszyngton:	A Standard Computable General Equilibrium (CGE) Model in GAMS. Washington:

Page	1st edition	corrected edition
191	McKenzie, & Lionel , W. (1959).	McKenzie, L.W. (1959).
191	McKenzie, & Lionel , W. (1981).	McKenzie, L.W. (1981).
191	McKenzie, & Lionel , W. (1987).	McKenzie, L.W. (1987).
191	McKenzie, L. (1954, April).	McKenzie, L.W. (1954).
192	Walras, L. (1874). Éléments d'économie politique pure, ou, Théorie de la richesse sociale. Lausanne: L. Corbaz & Cie.	delete
198	Steps of programming in GAMS: 4. Variable specifications 5. Equation specifications a. declaration b. algebraic structure specification 6. Model statement 7. Solve statement	Steps of programming in GAMS: 1. Variable specifications 2. Equation specifications a. declaration b. algebraic structure specification 3. Model statement 4. Solve statement
209	Golan A.	deleted
209	Edwin Jaynes	deleted
209	Stanley H. E.	deleted
209	Xavier Gabaix	deleted
209	Bottazzi G.	Bottazzi
209	Channing Arndt	Channing
209	Heckelei Thomas	Heckelei
209	Kantelhardt J. W.	Kantelhardt
209	Klump R.	Klump
209	Lévy P.P.	Lévy
209	Mandelbrot B.	Mandelbrot
209	Mounfield C.	Mounfield
209	Ormerod P.	Ormerod
209	Papageorgiou C.	Papageorgiou
209	Souma W.	Souma
209	Thierfelder Karen	Thierfelder
209	Weierstrass Karl	Weierstrass
209	Y. Ikeda	Ikeda